

# On one-step replica symmetry breaking in the Edwards-Anderson spin glass model

Gino Del Ferraro,<sup>1</sup> Chuang Wang,<sup>2,3</sup> Hai-Jun Zhou,<sup>2</sup> and Erik Aurell<sup>1,4,5</sup>

<sup>1</sup>*AlbaNova University Centre, KTH-Royal Institute of Technology  
Dept of Computational Biology, SE-106 91 Stockholm, Sweden*

<sup>2</sup>*State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics,  
Chinese Academy of Sciences, Beijing 100190, China*

<sup>3</sup>*School of Engineering and Applied Sciences, Harvard University,  
33 Oxford Street, Cambridge MA 02138 USA*

<sup>4</sup>*ACCESS Linnaeus Centre and Center for Quantum Materials,  
KTH-Royal Institute of Technology, SE-100 44 Stockholm, Sweden*

<sup>5</sup>*Depts of Applied Physics and Computer Science,  
Aalto University, FIN-00076 Aalto, Finland*

(Dated: July 26, 2016)

We consider a one-step replica symmetry breaking description of the Edwards-Anderson spin glass model in 2D. The ingredients of this description are a Kikuchi approximation to the free energy and a second-level statistical model built on the extremal points of the Kikuchi approximation, which are also fixed points of a Generalized Belief Propagation (GBP) scheme. We show that a generalized free energy can be constructed where these extremal points are exponentially weighted by their Kikuchi free energy and a Parisi parameter  $y$ , and that the Kikuchi approximation of this generalized free energy leads to second-level, one-step replica symmetry breaking (1RSB), GBP equations. We then proceed analogously to Bethe approximation case for tree-like graphs, where it has been shown that 1RSB Belief Propagation equations admit a Survey Propagation solution. We discuss when and how the one-step-replica symmetry breaking GBP equations that we obtain also allow a simpler class of solutions which can be interpreted as a class of Generalized Survey Propagation equations for the single instance graph case.

## I. INTRODUCTION

Low-temperature phases of systems with random interactions can be thought of as optimizing under a large set of highly interdependent and conflicting constraints. If a local optimum has been found it will typically be difficult to improve upon without adjusting many variables over a large domain, which is inherently difficult, for any optimization procedure. A physical process obeying detailed balance will furthermore care not directly about how different are two configurations but

rather about how high are the energy barriers between them; for low enough temperature transition times should follow an Arrhenius law,  $\tau_{transit} \sim \tau_{micro} e^{-\beta E_B}$ , where  $\tau_{micro}$  is a (microscopic) attempt time,  $E_B$  is a barrier height and  $\beta = \frac{1}{k_B T}$  is the inverse temperature. If barrier heights increase with system size  $N$  different such low-lying states will be essentially disconnected for very long times (exponential in  $N$ ), and the equilibrium Gibbs measure separates into a (weighted) sum over such states.

Let the above-mentioned (metastable) states be labeled by  $k$  and each characterized, over some long time, by an average energy  $E_k$  and a probability distribution  $P_k$  with entropy  $S_k$ . They can then be considered as having (metastable) free energy densities  $f_k = \frac{1}{N} (E_k - T S_k)$  and from these one can construct a second-order statistical mechanics built on a second-order partition function

$$\Xi(y) = \sum_k e^{-\beta y N f_k} \quad (1)$$

which is known as a one-step replica symmetry breaking (1RSB) scheme [1–4]. The Parisi parameter  $y$  in (1) plays the role of an inverse temperature and is conjugate to a free energy-like quantity (referred to as a generalized free energy) defined as  $G(y) = -\frac{1}{\beta y} \ln \Xi$ . In the large- $N$  limit the sum in (1) can be expected to be dominated by different subsets of states for different  $y$ , and knowing the corresponding maximizing distributions gives, in analogy with standard statistical mechanics, information on the system in the thermodynamic limit.

This program has been carried out with great success for dilute systems where the constraints are locally organized in tree-like structures. For such networks the states can be found as the fixed points of the iterative scheme known as Belief Propagation (BP) [5, 6], and the free energy of each state that enter in (1) is the corresponding Bethe approximation to the free energy. Furthermore, the computation of (1) can be carried out by an iterative scheme called Survey Propagation (SP), formally quite similar to BP [1, 7–10].

While there are several routes to SP theory, we will here only be concerned with a generalization of one presented in [1, 2] where the three steps are (i) an efficient computation of  $f_k$  using the Bethe approximation, (ii) an approximation of (1) by a Bethe approximation on the second order, and (iii) a reduction of this approximation to SP, which can also be efficiently computed. In spin glass language an analysis based on BP is referred to as the Replica Symmetric (RS) level, while an analysis based on SP is referred to a One-Step Replica Symmetry breaking (1RSB) level. We will use the term 1RSB to generally underline effects that follow from a sum over states as in (1) as opposed to considering only one state. As our focus here will be on systems in finite dimension and an analysis built on the Kikuchi approximation, we will also use the terms “Kikuchi approximation

on second level” and “second-level GBP” when discussing steps conceptually corresponding to (ii) above, and “Generalized Survey Propagation” (GSP) for (iii). We note that for the Kikuchi approximation and the analogy of Belief propagation (step (i) above) the term Generalized Belief Propagation (GBP) is well established [5, 11].

The rationale for considering a generalization of SP of this type which we call GSP is the simple fact that systems in finite dimension are not arranged in tree-like structures. The above sketched approach is therefore, for a physical system, only valid as a mean-field theory and is typically also presented as such. There exists a considerable literature as to whether these approaches apply to finite-dimensional systems *cf.* [12–15] – which we will not enter except to point out that the question of barriers, and their heights, is one of geometry and hence strongly dependent on dimensionality. Indeed, even for infinite-dimensional systems much less is known about the distribution of barrier heights than about the distributions of free energy minima, and (1) may therefore, as pertaining to dynamically separate states, even in idealized cases be somewhat of an ansatz.

The Kikuchi (or cluster) expansion was introduced to improve upon the mean-field estimates of thermodynamic quantities [16–18]. Algorithmically it can be turned into Generalized Belief Propagation (GBP) [11, 19, 20] which, though considerably more complicated than BP, can be used to estimate marginal probabilities on graphs with short loops and has applications to inference in lattice systems appearing in image processing problems [6]. The question then arises if the local minima of Kikuchi free energy, computed as fixed points of GBP, can be turned into a second-level statistical mechanics as in (1), and if that can be efficiently computed in analogy to SP. If that would be the case we could arrive at better systematic approximations to thermodynamics of finite-dimensional random systems and, perhaps conceptually more important, give a different kind of argument in favour of the spin glass approach in finite-dimensional system. This general issue has been already investigated in the recent literature for both the single instance and averaged case [21, 22] by using a different approach based on replicas which leads to a different class of equations. The differences between our and this approach will be discussed in Section V A.

In this paper we attempt to address the problem by directly writing a second-order partition function analogous to (1) for the Edwards-Anderson spin glass model in two dimensions, but using instead of Bethe approximation a Kikuchi expansion based on regions containing one, two or four spins (“vertices”, “rods” and “plaquettes”). This leads to a quite complex statistical model on second level, where the variables are GBP terms (“GBP messages”) obeying hard-core constraints (“GBP fixed point equations”) and weighted by terms  $e^{-\beta y N f_k}$  where the  $f_k$  are the Kikuchi free energies computed from the fixed points of GBP. We show how to perform a Kikuchi expansion for

this second-level model using regions isomorphic to the regions on the first level model, and hence also interpretable as (second-level) “vertices”, “rods” and “plaquettes”. In contrast to first level the amount of variables in each second-level region is however large, as will be described in some detail below. We also point out, as far as we know for the first time, that a Kikuchi expansion can be performed in many ways for such a second-level model, our choice is only one of the simplest. We are thus able to carry through the second step in the approach described above arriving at a definite – though complicated – set of equations which can be interpreted as 1RSB GBP equations which are not the same class of those obtained in [21] by the replica Cluster Variation Method (CVM), see discussion in Sec V A.

We then follow the same path which leads to SP, *i.e* assume the same ansatz, and discuss how the second-level GBP equations which we obtain here also admit a simpler class of solutions. These new equations can be interpreted as a generalized form of the Survey Propagation equations, in short a form of Generalized Survey propagation (GSP). We believe that this derivation is of interest for treating finite-dimensional systems in spin glasses as well as for computer science and satisfiability related problems. Although theoretically consistent, these equations however appear hard to solve computationally due to their large dimensionality and a product of different distributions as already pointed out in [21]. The purpose of the current work is thus to show an alternative approach to a generalized SP theory based on a new statistical model, in the same spirit as done in [1, 2] for the Survey Propagation approach.

The paper is structured as follows. In Section II we recall the Edwards-Anderson model and in Section III we describe a Kikuchi approximation for this model based on regions “vertices”, “rods” and “plaquettes”, leading to a Generalized Belief Propagation scheme. This presentation follows earlier contributions by two of us [23–25] to which we refer for some of the details. In Section IV we introduce and discuss a second-level GBP scheme and we present the 1RSB GBP equations at the plaquette region-graph level of the Kikuchi approximation. The differences between our approach and the the replica cluster variational approach of [21] are discussed in this section. In Section V, proceeding in analogy to the path which leads to the Survey Propagation algorithm, we show that this scheme does simplify similarly to the SP case to a class of generalized Survey Propagation equations. In Section VI we sum up and discuss our results. Four appendices integrate the main text and in the last of them we discuss, to our knowledge for the first time, meaning and computation of determinant factors which appear both in the SP [1, 2] and generalized SP theory, constructed by using a second-level region graph.

## II. THE EDWARDS-ANDERSON MODEL

The 2D Edwards-Anderson (EA) model [26] is defined on a finite dimensional square lattice by the Hamiltonian  $H = -\sum_{\langle ij \rangle} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i$ , where the first sum is over neighbouring dichotomic spin variables  $\sigma_i = \pm 1$ ,  $h_i$  is a local external field and the  $J_{ij}$  are quenched random variables. Using a factor graph representation with  $N$  variable nodes ( $i = 1, 2, \dots, N$ ) representing the spins and  $M$  factor nodes ( $a = 1, 2, \dots, M$ ) representing the interactions between the spins, the Hamiltonian can alternatively be written as:

$$H(\underline{\sigma}) = \sum_a^M E_a(\underline{\sigma}_{\partial a}) + \sum_i^N E_i(\sigma_i) \quad (2)$$

where  $\underline{\sigma} \equiv (\sigma_1, \sigma_2, \dots, \sigma_N)$  denotes a generic spin configuration, and by  $\underline{\sigma}_{\partial a}$  we mean  $\{\sigma_i | i \in \partial a\}$ , the set of variables which are involved in interaction  $a$ . In the case of a two dimensional lattice this is simply  $\underline{\sigma}_{\partial a} = \{\sigma_i, \sigma_j\}$ , but the more general notation will be convenient later. In the same representation the partition function can be written as:

$$Z(\beta) \equiv \sum_{\underline{\sigma}} e^{-\beta H(\underline{\sigma})} = \sum_{\underline{\sigma}} \prod_{i=1}^N \psi_i(\sigma_i) \prod_{a=1}^M \psi_a(\underline{\sigma}_{\partial a}) \quad (3)$$

where  $\beta = 1/k_B T$  is an inverse temperature and the terms

$$\psi_i(\sigma_i) \equiv e^{\beta h_i \sigma_i}, \quad \psi_a(\underline{\sigma}_{\partial a}) \equiv e^{\beta J_{ij}^{(a)} \sigma_i \sigma_j} \quad (4)$$

are called “potential functions” in BP terminology. That interaction is in the more general scheme denoted  $(a)$  hence for later convenience we include that label also for the interaction coefficient also  $J_{ij}^{(a)}$ . The equilibrium free energy is related to the partition function as

$$F(\beta) \equiv -\frac{1}{\beta} \ln Z(\beta). \quad (5)$$

The EA model has been extensively studied in the literature with the spin coupling constants  $J_{ij}$  following the  $\pm J$  bimodal distribution or the continuous Gaussian distribution with zero mean and variance  $J^2$ . If all the external magnetic fields  $h_i = 0$ , this model is in the spin glass phase when the inverse temperature  $\beta$  exceeds certain critical value  $\beta_c$ . The value of  $\beta_c$  is finite for systems in  $D \geq 3$  dimensions, while  $\beta_c = \infty$  for  $D = 2$ . The spin glass transition therefore occurs at zero temperature for two-dimensional systems [27–31].

### III. REGION GRAPH REPRESENTATION AND GENERALIZED BELIEF PROPAGATION EQUATIONS

In this section we introduce a region graph description of the 2-dimensional lattice and use the Cluster Variation Method (CVM) as presented by Kikuchi [16–18] to compute an approximation to the free energy and derive generalized belief propagation equations as presented in [20, 24, 25]. This method was invented to make corrections to the Bethe free energy and therefore, algorithmically, to the BP method for loopy graphs. Alternative approaches have been carried out using loop corrections, both for binary and non-binary variables, in [32–34] and more recently in [35].

A region-graph description of a lattice model defines a set  $R$  of regions that include some chosen basic clusters of nodes, their intersections, the intersection of their intersections, and so on. The basic idea underlying the CVM is to approximate the free energy as a sum of the free energies of these basic clusters of nodes, minus the free energy of over-counted cluster intersections, minus the free energy of the over-counted intersections of intersections, and so on [11]. Following the description in [20], we define a region average energy as  $U_\alpha(P_\alpha) = \sum_{\underline{\sigma}_\alpha} P_\alpha(\underline{\sigma}_\alpha) E_\alpha(\underline{\sigma}_\alpha)$ , where  $E_\alpha(\underline{\sigma}_\alpha)$  is the energy term of the region  $\alpha$  similar to those appearing in equations (2), and a region entropy  $S_\alpha(P_\alpha) = -k_B \sum_{\underline{\sigma}_\alpha} P_\alpha(\underline{\sigma}_\alpha) \log P_\alpha(\underline{\sigma}_\alpha)$ , in which  $\underline{\sigma}_\alpha \equiv \{\sigma_i | i \in \alpha\}$  denotes a microscopic configuration of the same region. The region probability  $P_\alpha$  is a variational parameter and has to be determined from minimization condition. Let us observe that, in cases where  $P_\alpha$  is the exact probability distribution of the region  $\alpha$ , the average energy correspond to the exact value whereas, differently, the region entropy remains an approximation [11]. The Kikuchi free energy functional is then introduced as the sum of all the region contributions:

$$F_K = \sum_{\alpha} c_{\alpha} \sum_{\underline{\sigma}_{\alpha}} P_{\alpha}(\underline{\sigma}_{\alpha}) \left[ E_{\alpha}(\underline{\sigma}_{\alpha}) + \beta^{-1} \log P_{\alpha}(\underline{\sigma}_{\alpha}) \right] \quad (6)$$

Above  $E_{\alpha}(\underline{\sigma}_{\alpha})$  is the energy term of the region  $\alpha$ , in the EA model, for instance,  $E_{\alpha}(\underline{\sigma}_{\alpha}) = \sum_{i \in \alpha} h_i \sigma_i + \sum_{a \in \alpha} J_{ij}^{(a)} \sigma_i \sigma_j$ . In (6) the region graph coefficients  $c_{\alpha}$  are counting number assigned to each region to ensure that each term only contributes once [25] and constructed recursively by:

$$c_{\gamma} = 1 - \sum_{\{\alpha: \alpha > \gamma\}} c_{\alpha}, \quad (7)$$

where with  $\alpha > \gamma$  we indicate that  $\alpha$  is an ancestor of  $\gamma$ , *i.e.* the regions  $\gamma$  is contained in the region  $\alpha$ . The index  $\alpha$  in (6) can in principle run over all the regions in the region graph although keeping all of them would be cumbersome. A Kikuchi approximation of the free energy at the  $R$ -level then fixes a set  $R$  of clusters made of maximal clusters and all their sub-clusters, and truncates the

expansion of the free energy (6) retaining only terms corresponding to clusters in  $R$ . The CVM aims to the minimization of the above free energy (6) under the constraints that the marginals  $P_\alpha$  are consistent. The saddle point equations derived from this minimization are the so called Generalized Belief Propagation equations. A detailed derivation of them, following this standard approach, is provided in Appendix A.

An alternative formulation of the Kikuchi approximation and a derivation of GBP equations can be obtained formulating a region-graph description in terms of factor graphs and messages, *i.e.* normalized probability distributions, between neighbouring regions. In what comes next we follow such a formulation as described in [25] in order to obtain consistency relations among regions which will be useful in the derivation of the 1RSB Generalized Belief Propagation equations (see Sec. IV). Within this description, generally a region  $\mu$  can be a parent of another region  $\nu$  if the set  $\nu$  is contained in the set  $\mu$ , and when this is the case we indicate the relation as  $\mu \rightarrow \nu$ . Hereafter we consider the squares or “plaquettes” to be the larger cluster  $R$  and therefore we deal with three types of clusters or regions: plaquettes, rods and vertices (see Fig. 1), although the description presented next is very general and not restricted to such maximal clusters. Each square contains four vertices and four interactions and is a parent of four rod regions; each rod region contains two vertices and one interaction and is a parent of two vertices; while each vertex region contains only one vertex and has no children. In Fig. 1, left panel, the parent-to-child relations  $\mu \rightarrow \nu$  are represented by black arrows for the regions highlighted in green, namely plaquette, rods and vertices. We recursively define an ancestor of a region to be a parent or the parent of an ancestor: a vertex thus has four parents (four rods) and four other ancestors (four plaquettes). Similarly we recursively define a descendant of a region to be a child or the child of a descendant. A plaquette thus has four children (four rods) and four other descendants (four vertices). A region graph  $R$ , within this description, is then a collection of all the regions and the specified parent-to-child relations (*i.e.*, arrows) between these regions.

Following a factor region graph description as in [24, 25], the partition function (3) can then be written as

$$Z(\beta) = \sum_{\underline{\sigma}} \prod_{\alpha \in R} \left[ \prod_{i \in \alpha} \psi_i(\sigma_i) \prod_{a \in \alpha} \psi_a(\underline{\sigma}_{\partial a}) \right]^{c_\alpha}. \quad (8)$$

where  $\alpha$  in principle ranges over all the possible regions. As noted in [25] it is possible to introduce sets of messages  $m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu)$ , *i.e.* normalized probability distributions, between parent-child regions in  $R$ , with the only constraint that  $\sum_{\sigma_\nu} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu) = 1$ . Indeed we observe that for each direct

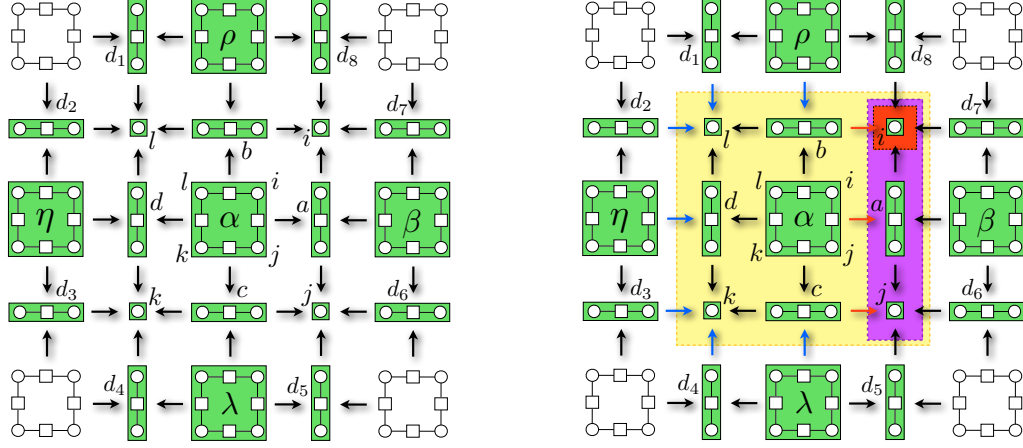


FIG. 1: Left panel: Region graph representation of a 2D EA model in terms of plaquettes, rods and vertices. Black arrows represent parent-to-child relations and are associated with messages from parents to child ( $m_{\mu \rightarrow \nu}$  in the main text). Right panel: The squared yellow region represents the union of the border ( $B_\alpha$ ) and interior ( $I_\alpha$ ) of the  $\alpha$  plaquette. The messages inside it belong to  $I_\alpha$  whereas the messages that cross the border belong to  $B_\alpha$ . The purple rod region represents the same object for the rod  $a$ . Blue and red arrows represent the messages involved in the GBP equation from plaquette  $\alpha$  to rod  $a$ , namely eq. (21) in the main text. The counting numbers for both figures are  $c_\alpha = 1$  for plaquettes,  $c_a = -1$  for rods and  $c_i = 1$  for vertexes.

edge  $\mu \rightarrow \nu$  the following relation holds

$$\sum_{\{\alpha : \mu \in B_\alpha, \nu \in I_\alpha\}} c_\alpha = \sum_{\alpha \geq \nu} c_\alpha - \sum_{\alpha \geq \mu} c_\alpha = 1 - 1 = 0, \quad (9)$$

where we have for each region  $\alpha$  defined  $I_\alpha \equiv \{\gamma : \gamma \leq \alpha\}$  to be the set formed by the region  $\alpha$  and all its descendants, and  $B_\alpha$ , the “boundary” of  $\alpha$ , *i.e.* to be the set of regions not belonging to  $I_\alpha$  but parental to at least one region in  $I_\alpha$  [25] (see Fig 1, right panel). Therefore the partition function (8) can be rewritten as:

$$Z(\beta) = \sum_{\underline{\sigma}} \prod_{\alpha \in R} \left[ \prod_{i \in \alpha} \psi_i(\sigma_i) \prod_{a \in \alpha} \psi_a(\underline{\sigma}_{\partial a}) \prod_{\{\mu \rightarrow \nu : \mu \in B_\alpha, \nu \in I_\alpha\}} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu) \right]^{c_\alpha} \quad (10)$$

indeed due to relation (9) each message appears in (10) is net total zero times. A Kikuchi approximation of this partition function corresponds to take into account only some chosen set of maximal clusters  $R$  in (10), their intersections, the intersection of their intersections, and so on. The Kikuchi approximation of the above partition function then reads as [24]

$$Z_K(\beta) = \prod_{\alpha \in R} \left[ \sum_{\underline{\sigma}_\alpha} \prod_{i \in \alpha} \psi_i(\sigma_i) \prod_{a \in \alpha} \psi_a(\underline{\sigma}_{\partial a}) \prod_{\{\mu \rightarrow \nu : \mu \in B_\alpha, \nu \in I_\alpha\}} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu) \right]^{c_\alpha} \quad (11)$$

where, differently from (10), here  $\alpha$  ranges over all regions up to a maximal cluster  $R$ , for instance restricted to plaquettes, rods and vertexes for the case discussed here. The corresponding Kikuchi



approximation of the free energy in terms of messages  $m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu)$ , reads as:

$$F_K(\beta) = -\beta^{-1} \sum_{\alpha \in R} c_\alpha \log \left[ \sum_{\underline{\sigma}_\alpha} \prod_{i \in \alpha} \psi_i(\sigma_i) \prod_{a \in \alpha} \psi_a(\underline{\sigma}_{\partial a}) \prod_{\{\mu \rightarrow \nu : \mu \in B_\alpha, \nu \in I_\alpha\}} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu) \right]. \quad (12)$$

Analogously to the standard derivation of the CVM, we then require this free energy to be stationary respect to a chosen set of probability functions  $\{m_{\mu \rightarrow \nu}\}$ , namely:

$$\frac{\delta F_K}{\delta m_{\mu \rightarrow \nu}} = 0, \quad \forall (\mu \rightarrow \nu) \in R. \quad (13)$$

As noted in [24, 25], this equation is satisfied if the consistency condition between parent-child regions of two marginal probability distributions is satisfied, namely

$$\sum_{\underline{\sigma}_\mu \setminus \underline{\sigma}_\nu} \omega_\mu(\underline{\sigma}_\mu) = \omega_\nu(\underline{\sigma}_\nu), \quad (14)$$

where  $\omega_\mu$  is a Boltzmann factor

$$\omega_\alpha(\underline{\sigma}_\alpha) \equiv \frac{1}{z_\alpha} \prod_{i \in \alpha} \psi_i(\sigma_i) \prod_{a \in \alpha} \psi_a(\underline{\sigma}_{\partial a}) \prod_{\{\mu \rightarrow \nu : \mu \in B_\alpha, \nu \in I_\alpha\}} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu) \quad (15)$$

called *belief* of the region  $\alpha$  in the message-passing language and  $z_\alpha$  is a normalization factor associated to the same region. Equation (14) ensures that the marginal probability distribution  $\omega_\mu(\underline{\sigma}_\mu)$  and  $\omega_\nu(\underline{\sigma}_\nu)$  of each parent-to-child pair come from the same joint probability distribution. The constraints on the marginals, together with (15), lead to the generalized belief-propagation equation on each directed edge  $\mu \rightarrow \nu$  of the region graph  $R$

$$\prod_{\substack{\alpha \rightarrow \gamma : \\ \{\alpha \in B_\nu \cap I_\mu, \gamma \in I_\nu\}}} m_{\alpha \rightarrow \gamma}(\underline{\sigma}_\gamma) = C \sum_{\underline{\sigma}_\mu \setminus \underline{\sigma}_\nu} \prod_{j \in \mu \setminus \nu} \psi_j(\sigma_j) \prod_{b \in \mu \setminus \nu} \psi_b(\underline{\sigma}_{\partial b}) \prod_{\substack{\{\eta \rightarrow \tau : \\ \eta \in B_\mu, \tau \in I_\mu \setminus I_\nu\}}} m_{\eta \rightarrow \tau}(\underline{\sigma}_\tau) \quad (16)$$

where  $C$  is a normalization constant determined by  $\sum_{\underline{\sigma}_\nu} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu) = 1$ .

### A. Plaquette-rod-vertex region graph description

As a concrete example and as a reference for the following sections, we explicitly give the Kikuchi partition function (free energy) and the corresponding generalized belief propagation equations for a region graph description of the 2D Edward-Anderson model. Letting now Greek letter  $\alpha$  stand

for plaquettes,  $a$  for rods and  $i$  for vertices and introducing auxiliary functions

$$Z_\alpha = \sum_{\underline{\sigma}_\alpha} \prod_{i \in \alpha} \psi_i(\sigma_i) \prod_{a \in \alpha} \psi_a(\underline{\sigma}_{\partial a}) \prod_{\{\mu \rightarrow \nu \mid \mu \in B_\alpha, \nu \in I_\alpha\}} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu), \quad (17)$$

$$Z_a = \sum_{\underline{\sigma}_a} \psi_a(\underline{\sigma}_{\partial a}) \prod_{i \in a} \psi_i(\sigma_i) \prod_{\{\mu \rightarrow \nu \mid \mu \in B_a, \nu \in I_a\}} m_{\mu \rightarrow \nu}(\underline{\sigma}_\nu), \quad (18)$$

$$Z_i = \sum_{\sigma_i} \psi_i(\sigma_i) \prod_{\{\mu \rightarrow i \mid \mu \in B_i\}} m_{\mu \rightarrow i}(\sigma_i), \quad (19)$$

the Kikuchi partition function (11) can be written compactly as

$$Z_K = e^{-\beta F_K} = \prod_{\alpha \in R} Z_\alpha \prod_{a \in \alpha} Z_a^{-1} \prod_{i \in a} Z_i \quad (20)$$

where  $F_K$  refers to the Kikuchi free energy (12). As messages go from parents to children, and as in the chosen region graph we have only parents (plaquettes), children (rods) and grand-children (vertices), we have only two instances of the general equation (16). With reference to Fig. 1, right panel, for the labeling, the first is plaquettes-to-rods and reads

$$\begin{aligned} m_{\alpha \rightarrow a}(\sigma_i, \sigma_j) m_{b \rightarrow i}(\sigma_i) m_{c \rightarrow j}(\sigma_j) &= \frac{1}{Z_{\alpha \rightarrow a}} \sum_{\sigma_l, \sigma_k} \psi_l(\sigma_l) \psi_k(\sigma_k) \psi_b(\sigma_l, \sigma_i) \psi_d(\sigma_l, \sigma_k) \psi_c(\sigma_k, \sigma_j) \\ &\quad \times m_{\rho \rightarrow b}(\sigma_i, \sigma_l) m_{d_1 \rightarrow l}(\sigma_l) m_{d_2 \rightarrow l}(\sigma_l) m_{\eta \rightarrow d}(\sigma_l, \sigma_k) \\ &\quad \times m_{d_3 \rightarrow k}(\sigma_k) m_{d_4 \rightarrow k}(\sigma_k) m_{\lambda \rightarrow c}(\sigma_k, \sigma_j) \end{aligned} \quad (21)$$

where  $Z_{\alpha \rightarrow a}$  is a normalization factor. The messages which appear on the LHS of (21) are those depicted with red arrows in Fig 1 (right panel) whereas those on the RHS are depicted as blue arrows. Equation (21) can be written compactly, in a form we will use later, as

$$m_{\alpha \rightarrow a}(\underline{\sigma}_a) = \mathcal{F}_{\alpha \rightarrow a}(\{m_{\eta \rightarrow \tau}\}_{[\eta \in B_\alpha, \tau \in I_\alpha \setminus I_a]}, \{m_{\beta \rightarrow \gamma}\}_{[\beta \in (B_a \cap I_\alpha) \setminus \alpha, \gamma \in I_a \setminus a]}) \quad (22)$$

The second instance is the rods-to-vertices equation and reads

$$m_{a \rightarrow i}(\sigma_i) = \frac{1}{Z_{a \rightarrow i}} \sum_{\sigma_j} \psi_j(\sigma_j) \psi_a(\sigma_i, \sigma_j) m_{\alpha \rightarrow a}(\sigma_i, \sigma_j) m_{\beta \rightarrow a}(\sigma_i, \sigma_j) m_{c \rightarrow j}(\sigma_j) m_{d_5 \rightarrow j}(\sigma_j) m_{d_6 \rightarrow j}(\sigma_j) \quad (23)$$

where  $Z_{a \rightarrow i}$  is a normalization factor. Referring to Fig. 2 (right panel), the messages on the RHS of (23) are those depicted in black whereas the message on the LHS is depicted in white. Similarly to (21), equation (23) can also be written in a compact form as

$$m_{a \rightarrow i}(\sigma_i) = \mathcal{F}_{a \rightarrow i}(\{m_{\eta \rightarrow \tau}\}_{[\eta \in B_a, \tau \in I_a \setminus I_i]}) \quad (24)$$

The GBP equations (21) and (23) have a somewhat abstract flavor. In practice they should be considered as iterative equations for the real parameters describing the various probability distributions. For completeness we give such a more detailed description in Appendix B below. The region graph shown in Fig. 1 is redundant in the sense that there are two directed paths from a plaquette region  $\alpha$  to each of its four vertex child regions [24]. Because of this redundancy the LHS of the GBP equation (21) is a product of three messages, which leads to practical complications in updating the plaquette-to-rod message  $m_{\alpha \rightarrow a}(\sigma_i, \sigma_j)$ . In general we regard a region graph as a non-redundant region graph if there is at most one directed path from any region to any another region, otherwise the region graph is regarded as redundant [24]. In the case of a non-redundant region graph the GBP equation can be much simplified and the LHS of each GBP equation contains only one message (see [24] for detailed discussions). Although redundant region graphs bring in computational complications, the main reason to prefer redundant region graphs over non-redundant ones is that, in a redundant region graph, more consistency constraints are enforced on the parent-to-child messages and consequently the results are more accurate.

## B. Gauge invariance of the GBP equations

When the GBP equations are written to be iterated, for convenience, the messages are parametrized in terms of fields by using an exponential representation. In principle the messages can be left not-normalized therefore the following parametrization does not take into account their normalization which, when and if needed, can however be enforced at any time. Messages  $m_{a \rightarrow i}(\sigma_i)$  are written as  $e^{\beta u_{a \rightarrow i} \sigma_i}$  where  $u_{a \rightarrow i}$  is a real number known as the cavity bias on spin  $i$  from interaction  $a$ . Similarly, messages  $m_{\alpha \rightarrow a}(\sigma_i, \sigma_j)$  are written as  $e^{\beta (u_{\alpha \rightarrow a}^{(i)} \sigma_i + u_{\alpha \rightarrow a}^{(j)} \sigma_j + U_{\alpha \rightarrow a} \sigma_i \sigma_j)}$  where the new quantity  $U_{\alpha \rightarrow a}$  parametrizes cooperative interactions. The GBP equations (21) and (23) are then to be understood as relations between parameters of the types  $u_{a \rightarrow i}^{(i)}$ ,  $u_{\alpha \rightarrow a}^{(i)}$  and  $U_{\alpha \rightarrow a}$ . In Appendix B a derivation and the explicit expression for the GBP equations in terms of fields is reported. Here we just show their functional dependence which is the main thing for what follows. The link-to-vertex equation in term of fields reads as [36]

$$u_{a \rightarrow i} = \mathcal{L}_{a \rightarrow i}(u_{\alpha \rightarrow a}^{(i)} + u_{\beta \rightarrow a}^{(i)}, J_{ij}^{(a)} + U_{\alpha \rightarrow a} + U_{\beta \rightarrow a}, h_j + u_{\alpha \rightarrow a}^{(j)} + u_{\beta \rightarrow a}^{(j)} + u_{c \rightarrow j} + u_{d_5 \rightarrow j} + u_{d_6 \rightarrow j}) \quad (25)$$

(see Appendix B for the explicit formula). In Figure 2, right panel, the message  $u_{a \rightarrow i}$  on the LHS of (25) is illustrated as a white arrow whereas all the messages which appear on the RHS are illustrated as black arrows. The equations for the fields involved in the plaquette-to-link equations,

using a similar notation as in [36], are given by

$$U_{\alpha \rightarrow a} = \mathcal{F}_{\alpha \rightarrow a}(\vec{U}, \vec{u}) = \frac{1}{4\beta} \log \left[ \frac{K(1, 1)K(-1, -1)}{K(1, -1)K(-1, 1)} \right], \quad (26)$$

$$u_{\alpha \rightarrow a}^{(i)} = \mathcal{G}_{\alpha \rightarrow a}^{(i)}(\vec{U}, \vec{u}) = u_{\rho \rightarrow b}^{(i)} - u_{b \rightarrow i} + \frac{1}{4\beta} \log \left[ \frac{K(1, 1)K(1, -1)}{K(-1, 1)K(-1, -1)} \right], \quad (27)$$

$$u_{\alpha \rightarrow a}^{(j)} = \mathcal{G}_{\alpha \rightarrow a}^{(j)}(\vec{U}, \vec{u}) = u_{\lambda \rightarrow c}^{(j)} - u_{c \rightarrow j} + \frac{1}{4\beta} \log \left[ \frac{K(1, 1)K(-1, 1)}{K(1, -1)K(-1, -1)} \right], \quad (28)$$

where  $K(\sigma_i, \sigma_j)$  is defined in the Appendix B and only depends on the messages in  $B(\alpha)$ , which we here call “external” field-messages, *i.e.* for the  $\alpha \rightarrow a$  equations only those illustrated as blue arrows entering the yellow square region in Figure 2, left panel. Therefore from (26) we observe that the field-messages of the kind  $U_{\alpha \rightarrow a}$  do not depend on any messages in  $I(\alpha)$  that we call “internal” message, whereas from (27) and (28) we note that the dependence on internal messages appears linearly only through the messages of the kind  $u_{b \rightarrow i}$  and  $u_{c \rightarrow j}$  respectively.

In [36] the authors point out that parent-to-child GBP algorithms, like the one we use here, posses a gauge invariance on the field-message values. In fact, GBP equations for such algorithms admit a freedom in the choice of the fields that has no effect on the fixed point solutions. Therefore, following [36], every set of values of the fields which satisfies a fixed point can be changed by an arbitrary constant  $\delta$  as

$$\begin{aligned} u_{a \rightarrow i} &\rightarrow u_{a \rightarrow i} + \delta, & u_{\alpha \rightarrow a}^{(i)} &\rightarrow u_{\alpha \rightarrow a}^{(i)} + \delta, \\ u_{b \rightarrow i} &\rightarrow u_{b \rightarrow i} - \delta, & u_{\alpha \rightarrow b}^{(i)} &\rightarrow u_{\alpha \rightarrow b}^{(i)} - \delta, \end{aligned} \quad (29)$$

and the resulting set of values is still a solution of the GBP equations. This gauge invariance can be fixed in several ways, one of the simplest is to set to zero one of the four  $u$ -field in the equation (29). This gauge fixing method will be used in the following sections when building a second-level statistical model based on the GBP fixed point solutions.

#### IV. ONE-STEP REPLICA SYMMETRY BREAKING GBP EQUATIONS

We here take up the approach discussed in the Introduction and consider a second-order statistical model built on a second level (1RSB) partition function (1). In the picture we want to consider here, the Gibbs measure is, in principle, not anymore a pure state but rather decomposes into a convex linear combination of pure states [1–4]. We assume that in each such state labeled by  $k$  the GBP equations (16) are satisfied and we denote with  $\{m_{\mu \rightarrow \nu}^{(k)}\}$  the set of their solution, for every  $\mu \rightarrow \nu \in R$ . This assumption is supported by the recent results in [37] where it is shown that metastable states found with Monte Carlo algorithm are linked to fixed points of the Cluster

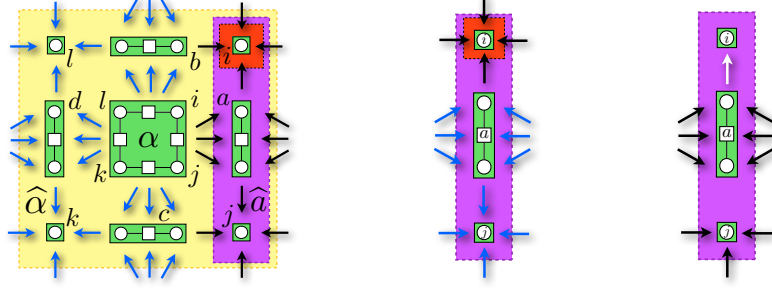


FIG. 2: Left panel: First and second level region graph description overlapped for the 2D EA model. Messages integrated over on the LHS of (37) are illustrated with blue arrows whereas those on the RHS are pictured with black arrows. Central panel: same kind of illustration for equation (38). Right panel: pictorial representation of GBP rod-to-vertex equation (23). The message on the LHS is pictured with a white arrow whereas black arrows illustrate the messages on the RHS.

Variational Method. A Kikuchi approximation of the generalized free energy (1) is then given by replacing the exact free energy of every state by the Kikuchi free energy computed at the same state, *i.e.*  $Nf_k \rightarrow F_K\{m_{\mu \rightarrow \nu}^{(k)}\}$  in (1). The resulting second-order partition function is defined as a sum over fixed points of GBP weighted by their respective Kikuchi free energy:

$$\Xi(y) = \sum_k e^{-\beta y N f_k} \simeq \sum_{k \in E} e^{-\beta y F_K\{m_{\mu \rightarrow \nu}^{(k)}\}} \equiv \Xi_K(y), \quad (30)$$

where  $E$  denotes the set of GBP fixed points and with  $\Xi_K$  we define the Kikuchi approximation of the replicated or second order partition function. By standard folklore we expect either GBP to have only one fixed point, in which case the sum and the additional parameter  $y$  in (30) gives no new information, or to have many fixed points, in which case their relative contributions to the sum are controlled by  $y$ . The messages  $m_{\mu \rightarrow \nu}$  are considered variables in a statistical model, and the restrictions that the messages are fixed points of GBP are interpreted as hard-core constraints

$$\Xi_K = \sum_{k \in E} \int \mathcal{D}m_{\mu \rightarrow \nu} e^{-\beta y F_K\{m_{\mu \rightarrow \nu}\}} \prod_{\langle \mu \rightarrow \nu \rangle} \delta[m_{\mu \rightarrow \nu} - m_{\mu \rightarrow \nu}^{(k)}] \quad (31)$$

where on each pair of connected regions  $\langle \mu \rightarrow \nu \rangle$ , plaquette to link and link to vertex, we have a message  $m_{\mu \rightarrow \nu}$  and the delta functions enforce that these variables correspond to a given fixed point, labeled by  $(k)$ , of the GBP equations.

Actually we cannot however sum over an enumeration of the fixed points of GBP, that would not be any simplification of the problem. Instead we must include the hard-core constraints inside

the integration as

$$\Xi_K(y) = \int \mathcal{D}m_{\mu \rightarrow \nu} e^{-\beta y F_K\{m_{\mu \rightarrow \nu}\}} \prod_{\langle \mu \rightarrow \nu \rangle} \delta[m_{\mu \rightarrow \nu} - \mathcal{F}_{\mu \rightarrow \nu}(\underline{\hat{m}})] \quad (32)$$

where  $\underline{\hat{m}}$  indicates all the dependencies of the function  $\mathcal{F}$  from the messages involved in the GBP equations. By postulating (32) we have ignored Jacobians from integration of the delta functions. If we insist that  $\Xi(y)$  in (30) is only a sum over fixed points, with no pre-factors, we should more properly have

$$\Xi_K(y) = \int \mathcal{D}m_{\mu \rightarrow \nu} e^{-\beta y F_K\{m_{\mu \rightarrow \nu}\}} \left| \det \left[ \frac{\partial(m_{\mu \rightarrow \nu} - \mathcal{F}_{\mu \rightarrow \nu})}{\partial m_{\mu' \rightarrow \nu'}} \right] \right| \prod_{\langle \mu \rightarrow \nu \rangle} \delta[m_{\mu \rightarrow \nu} - \mathcal{F}_{\mu \rightarrow \nu}(\underline{\hat{m}})] \quad (33)$$

where the determinant is of a large matrix - coupling the first-level variables (GBP messages) as they are connected by the hard-core constraints (GBP update equations) - computed at the fixed point of the GBP equations. This value is independent on the Parisi parameter  $y$  and therefore contributes only to the entropic part of re-weighting term  $e^{-\beta y F_K\{m_{\mu \rightarrow \nu}\}}$  in equation (33). For tractability we will from here neglect this term, to discuss it again in Section VI and Appendix D. We just mention that a similar determinant appears also for a tree-like topology and in that case it is similarly neglected as discussed in [1, 2]. Let us also observe that the above integrals in (32) and (33) converge only if the gauge invariance of the GBP equation is fixed (see Section III B), otherwise a volume element of the parametrization invariance appears. We assume here that this is the case and we specify in more details how to fix the gauge invariance at the second-level statistical model in Section IV C. Using (20) and (32) we can then write

$$\Xi_K = \sum_{k \in E} e^{-\beta y F_K\{m_{\mu \rightarrow \nu}^{(k)}\}} = \int \mathcal{D}m_{\mu \rightarrow \nu} \prod_{\alpha \in R} Z_{\alpha}^y \prod_{a \in R} Z_a^{-y} \prod_{i \in R} Z_i^y \left[ \prod_{\langle \mu \rightarrow \nu \rangle} \delta[m_{\mu \rightarrow \nu} - \mathcal{F}_{\mu \rightarrow \nu}(\underline{\hat{m}})] \right] \quad (34)$$

where we recall that  $Z_{\alpha}$ ,  $Z_a$  and  $Z_i$  are functions of the messages as given in (17), (18) and (19). We will further show below that each  $\delta$ -function in (34) can be assigned to one or more regions on a higher-level graph. We therefore interpret equation (34) as the partition function of a new model where the messages represent the new variables, and the  $Z$ 's together with an appropriate set of  $\delta$ -functions are the analogues on the second level of GBP potential functions on the first level. We note that these second-level potential functions contain both soft constraints in the  $Z$ 's and hard constraints in the  $\delta$ -functions.

### A. Higher-level region graph and second-level GBP

In the previous section we introduced a second-level graphical model by introducing the new partition function (34). The next step towards 1RSB GBP equations is to observe that the weighting

terms  $w_k = \exp(-\beta y F_K\{m_{\mu \rightarrow \nu}^{(k)}\})/\Xi_K$  induce a distribution of the messages and that, furthermore, this distribution can be represented as a graphical model. In this new second-level region-graph model the new variables are messages and the factor nodes are new potential functions or constraints. Each factor node is connected to all the variable nodes on which it depends *i.e.* for a potential functions to all the arguments of respectively  $Z_\alpha$ ,  $Z_a$  and  $Z_i$ , and for a constraint  $\delta[U_{\mu \rightarrow \nu} - \mathcal{F}_{\mu \rightarrow \nu}(\hat{U}, \hat{u})]$  or  $\delta[u_{\mu \rightarrow \nu} - \mathcal{F}_{\mu \rightarrow \nu}(\hat{U}, \hat{u})]$  both to  $U_{\mu \rightarrow \nu}$  or  $u_{\mu \rightarrow \nu}$  and to all the arguments of  $\mathcal{F}_{\mu \rightarrow \nu}$ . This auxiliary model is a graph expansion of the original model, as discussed for the case of BP in [1, 2].

The second step is to define new regions in the auxiliary graph and assign the second-level potential functions and constraints to these regions. How to do so is not uniquely defined, no more than in the standard CVM. We here choose a region graph for the auxiliary second-level graphical model which is isomorphic to the region graph for the first-level graphical model. This means that plaquettes, rods and vertices on the first level will be in one-to-one correspondence with regions on the second level which we call by the same names. To a plaquette  $\alpha$  on the first-level graph corresponds a second-level region denoted by  $\hat{\alpha}$  which contains (i) all messages going to descendants of  $\alpha$  from the outside or from the inside *i.e.*  $B_\alpha \cup I_\alpha$ ; (ii) the delta functions enforcing that the messages in  $I_\alpha$  satisfy GBP; (iii) the potential function  $Z_\alpha^y$ . Similarly, to a rod  $a$  on the first-level graph corresponds a second-level region denoted by  $\hat{a}$  which contains all messages in  $B_a \cup I_a$ , the two delta functions enforcing that the messages in  $I_a$  satisfy GBP, and  $Z_a^y$ . To a vertex  $i$  corresponds finally  $\hat{i}$  which contains the four messages in  $B_i$ , and  $Z_i^y$ . The procedure is illustrated in Fig 3.

The third step is to carry out CVM on (34) with the regions defined in Fig. 3 and define consistency conditions for marginals between parent-to-child regions similarly to those in (14) for the first-level region graph. The second-level CVM is built on messages in the second-level region graph, which are functions of the messages  $\vec{m}$  or, rather, of the fields used to parametrize the messages in the first-level region graph and which we here denote as  $Q(\vec{U}, \vec{u})$  (probabilities of probabilities) illustrated by black arrows in Fig 3, right panel. Every one of these (second-level) messages goes from a parent to a child region and depends on all the (first-level) messages, *i.e.* message-fields, in the child region. When there is a direct plaquette-to-rod link  $\hat{\alpha} \rightarrow \hat{a}$  on the second-level region graph, the  $Q_{\hat{\alpha} \rightarrow \hat{a}}$  message depends on all the fields in the region  $\hat{a}$ , that is all the fields in  $B_a \cup I_a$  in the original graph and so depends on 14 (first-level) field-messages. When there is a direct rod-to-vertex link  $\hat{a} \rightarrow \hat{i}$  the  $Q_{\hat{a} \rightarrow \hat{i}}$  message depends on all the fields in the region  $\hat{i}$ , corresponding to the region  $B_i$  in the original graph and so includes 4 fields (see Fig 3 for an

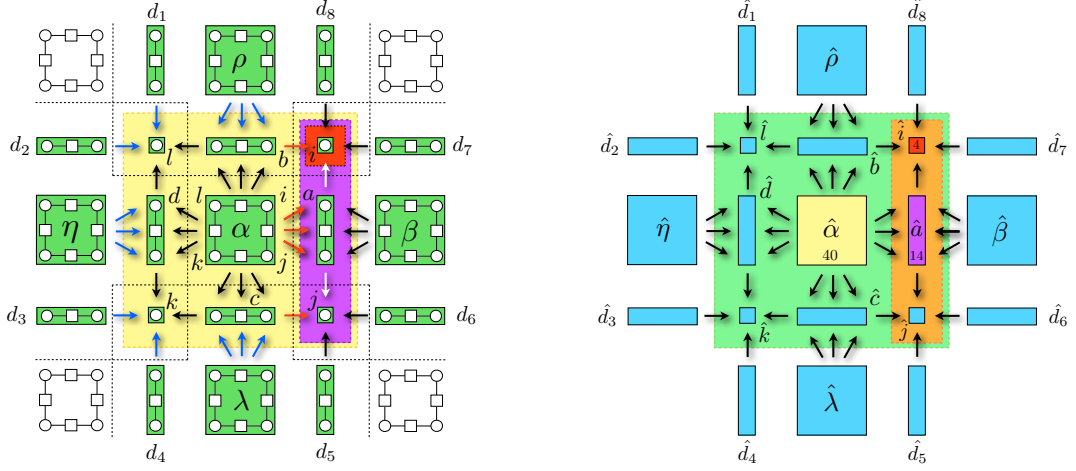


FIG. 3: Left panel: Region graph representation of a 2D EA model in terms of plaquettes, rods and vertices. The border and interior of the plaquette region  $\alpha$  is pictured in yellow and becomes the region  $\hat{\alpha}$  in the auxiliary model (yellow plaquette in the right panel). The border and interior of the rod region  $a$  is pictured in purple and becomes the new rod region  $\hat{a}$  (right panel) whereas the border of the vertex region  $i$ , which is displayed in red, becomes the new vertex region  $\hat{i}$  in the auxiliary model. Right panel: the auxiliary graph in terms of new regions. Numbers inside the regions are a memo of the number of  $\vec{U}, \vec{u}$  field-messages contained in each region. The arrows represent the new messages in the auxiliary model, indicated as  $Q(\vec{U}, \vec{u})$  in the text.

illustration).

The CVM and GBP construction at the first level introduce two types of  $\vec{m}$  first-level messages, from plaquette-to-rod and from rod-to-vertex, which can be parametrized using  $u$ -fields as shown in III B. Therefore on the second-level region, by a one-to-one correspondence with the first-level graph, there also exist two different kinds of  $Q(\vec{U}, \vec{u})$  messages. More explicitly, we represent these second-level messages in the auxiliary graph using a different notation for each one of them:

$$Q_{\hat{p} \rightarrow \hat{r}}(\vec{U}_{p \rightarrow r}^{(2)}, \vec{u}_{p \rightarrow r}^{(v)(4)}, \vec{u}_{r \rightarrow v}^{(8)}) \quad \text{from plaquettes to rods,} \quad (35)$$

$$q_{\hat{r} \rightarrow \hat{v}}(\vec{u}_{r \rightarrow v}^{(4)}) \quad \text{from rods to vertices,} \quad (36)$$

where superscript numbers refers to how many messages of the kind  $\mu \rightarrow \nu$  are contained in the argument of  $Q$ . That is, the notation  $U_{p \rightarrow r}^{(l)}$  means  $l$  field-messages of the kind plaquette-to-rod,  $\vec{u}_{p \rightarrow r}^{(v)(l)}$  means  $l$  field-messages of the kind plaquette-to-rod of the vertex type, whereas  $\vec{u}_{r \rightarrow v}^{(l)}$  means  $l$  messages of the kind rod-to-vertex, according to notation introduced in Section III B. In the following, for the sake of clarity, we omit the labels and the numbers of the fields  $\vec{u}$  when they are arguments of the functions  $Q$  and  $q$ , *i.e.* we simply write  $Q_{\hat{p} \rightarrow \hat{r}}(\vec{U}, \vec{u})$  and  $q_{\hat{r} \rightarrow \hat{v}}(\vec{u})$ .

We can now look at the second-level region graph as a new auxiliary region graph on an higher



level, which contains new messages (probabilities of probabilities), new potential functions and hard-core constraints. As for the first level-region graph (see eq. (14)), the new marginal probabilities, or beliefs, defined on the second-level region graph, which we here indicate as  $W_{\hat{\alpha}}$ ,  $W_{\hat{a}}$  and  $W_{\hat{i}}$  for new plaquette, rod and vertex respectively, must satisfy probability consistency equation as

$$\int_{\hat{\alpha} \setminus \hat{a}} d\vec{U}^{(6)} d\vec{u}^{(20)} W_{\hat{\alpha}}(\vec{U}^{(8)}, \vec{u}^{(32)}) = W_{\hat{a}}(\vec{U}^{(2)}, \vec{u}^{(12)}) \quad \text{for plaquette } \hat{\alpha} \text{ and rod } \hat{a}, \quad (37)$$

$$\int_{\hat{a} \setminus \hat{i}} d\vec{U}^{(2)} d\vec{u}^{(8)} W_{\hat{a}}(\vec{U}^{(2)}, \vec{u}^{(12)}) = W_{\hat{i}}(\vec{u}^{(4)}) \quad \text{for rod } \hat{a} \text{ and vertex } \hat{i}. \quad (38)$$

Above the notation  $\hat{\alpha} \setminus \hat{a}$  means that the integral is taken over all the messages belonging to  $\hat{\alpha}$  except those which are common with  $\hat{a}$  (see Fig. 2, left panel) (and similarly for  $\hat{a} \setminus \hat{i}$ , see Fig. 2, central panel). Analogously to the marginals (15) in the first-level region graph, the second-level beliefs are defined through the new messages  $Q$  and  $q$ , new potential functions  $Z$  and hard-constraints included in each second-level region. Since equations (37) and (38) are the core of our approach to second-level GBP we will write them down explicitly in the following two subsections.

### B. Consistency equation for rods and vertices

In analogy with the consistency equations (14), we here write explicitly the consistency equations for the second-level region graph illustrated in Fig. 3. As mentioned, we interpret the second-level graph as a new auxiliary graph having new potential functions. These are given by the partition functions  $Z$  re-weighted by the Parisi parameter  $y$ , new messages between parent-to-child regions given by the probabilities of probabilities  $Q$  and  $q$  and, in addition to the original region graph, contains new constraints which enforce the new variables  $\vec{U}$  and  $\vec{u}$  of the second-level graph to satisfy the GBP conditions on the first-level region graph.

The new marginals can then be deduced by analogy with the first-level marginals. The  $Q$  and  $q$  messages involved in the marginal of a region are all those from parent-to-child region  $\hat{\mu} \rightarrow \hat{\nu}$  where  $\hat{\mu}$  belongs to the border and  $\hat{\nu}$  belongs to the interior of the child region. The potential function for each second-level region must count for all the energetic interactions of that region and therefore it is given by the partition function of the same region which contains, indeed, all the energetic terms. Finally the hard-core constraints should enforce all the GBP conditions for the messages internal to the region and therefore they have to be all the delta functions which belong to the interior of the region considered.

With this procedure as a tool, using a compact notation, the local consistency equation (38)

can be written as follows:

$$\frac{1}{\Xi_{\hat{a}}} \int_{\hat{a} \setminus \hat{i}} d\vec{U} d\vec{u} Z_a^y(\vec{U}, \vec{u}) \delta_{a \rightarrow i} \delta_{a \rightarrow j} \prod^{(2)} Q_{\hat{\mu} \rightarrow \hat{\nu}}(\vec{U}, \vec{u}) \prod^{(6)} q_{\hat{\mu} \rightarrow \hat{\nu}}(\vec{u}) = \frac{1}{\Xi_{\hat{i}}} Z_i^y(\vec{u}) \prod^{(4)} q_{\hat{b} \rightarrow \hat{i}}(\vec{u}) \quad (39)$$

where  $\Xi_{\hat{a}}$  and  $\Xi_{\hat{i}}$  are normalization factors and the superscript on the products indicate how many  $Q$  or  $q$  products we have on each side. The delta functions in the equation above enforce the GBP equations (25) for rods-to-vertex messages and are interpreted as hard-core constraints; for short we use the notation  $\delta_{\mu \rightarrow \nu} = \delta[u_{\mu \rightarrow \nu} - \mathcal{L}_{\mu \rightarrow \nu}(\hat{U}, \hat{u})]$ , where the function  $\mathcal{L}_{\mu \rightarrow \nu}$  is defined in (25) or explicitly in (B5). The messages constrained by these two deltas are pictorially represented as white arrows in the purple rod region of Figure 3, left panel. The  $Z$ 's are potential functions referred to the region considered and the Parisi parameter  $y$  plays the role of an inverse temperature. The different kind of  $Q$ 's on the LHS are all those from the parent-to-child region  $\hat{\mu} \rightarrow \hat{\nu}$  with  $\hat{\mu} \in B(\hat{a})$  and  $\hat{\nu} \in I(\hat{a})$  and the products are on these same variables. Pictorially, they are represented by all the arrows crossing the perimeter of the orange rod region of Figure 3, right panel. Analogously, all the  $Q$ 's and the products on the RHS are all those from  $\hat{\mu} \rightarrow \hat{i}$  where  $\hat{\mu} \in B(\hat{i})$  therefore, pictorially, the four black arrows pointing towards the red region  $\hat{i}$  in the same figure.

Similarly to what pointed out for equations (32) and (33), the introduction of delta functions with functional dependence on the GBP equations above would, in principle, involve a determinant on the LHS of (39). This determinant is of the type  $|\det(\mathbb{1} - \Delta \mathcal{L})|$  where  $\Delta \mathcal{L}$  is the matrix containing the derivatives of the GBP rod-to-vertex function  $\mathcal{L}(\vec{U}, \vec{u})$ , respect to all the fields  $\vec{U}$  and  $\vec{u}$  belonging to the rod region  $\hat{a}$ , computed at the fixed point. As shown in Appendix D, it turns out that this determinant is equal to one and therefore brings no correction to the above equation.

Writing more explicitly (39) for the rod  $\hat{a}$  and the vertex  $\hat{i}$  on the expanded graph, with reference to Fig. 3 for labelling, we have

$$\begin{aligned} \frac{1}{\Xi_{\hat{a}}} \int_{\hat{a} \setminus \hat{i}} d\vec{U} d\vec{u} Z_a^y(\vec{U}, \vec{u}) \delta_{a \rightarrow i} \delta_{a \rightarrow j} Q_{\hat{\alpha} \rightarrow \hat{a}}(\vec{U}, \vec{u}) Q_{\hat{\beta} \rightarrow \hat{a}}(\vec{U}, \vec{u}) \underbrace{q_{\hat{b} \rightarrow \hat{i}}(\vec{u}) q_{\hat{d}_8 \rightarrow \hat{i}}(\vec{u}) q_{\hat{d}_7 \rightarrow \hat{i}}(\vec{u})}_{\text{dashed line}} \\ \times \underbrace{q_{\hat{c} \rightarrow \hat{j}}(\vec{u}) q_{\hat{d}_5 \rightarrow \hat{j}}(\vec{u}) q_{\hat{d}_6 \rightarrow \hat{j}}(\vec{u})}_{\text{dashed line}} = \frac{1}{\Xi_{\hat{i}}} Z_i^y(\vec{u}) q_{\hat{a} \rightarrow \hat{i}}(\vec{u}) \underbrace{q_{\hat{b} \rightarrow \hat{i}}(\vec{u}) q_{\hat{d}_8 \rightarrow \hat{i}}(\vec{u}) q_{\hat{d}_7 \rightarrow \hat{i}}(\vec{u})}_{\text{dashed line}} \end{aligned} \quad (40)$$

where the common terms on the left and right-hand-side have been underlined by a dashed line. Let us observe that the integration does not involve these underlined terms. Indeed these terms only depend on the  $\vec{u}$  field-messages which belong to the region  $\hat{i}$  whereas the integration is over the  $\vec{u}$  in the set  $\hat{a} \setminus \hat{i}$ . Referring to Figure 2, central panel, where the  $\vec{u}$  are illustrated by arrows, the integration only involves the blue  $\vec{u}$ 's in the purple region  $\hat{a}$  and does not include the black fields

shared with the red region  $\hat{i}$ . Therefore since the underlined  $q$ 's never depend on the variable of integration they can be canceled on both sides of the equation. The result reads

$$q_{\hat{a} \rightarrow \hat{i}}(\vec{u}) = \frac{\Xi_{\hat{i}}}{\Xi_{\hat{a}}} \int_{\hat{a} \setminus \hat{i}} d\vec{U} d\vec{u} Z_{a \rightarrow i}^y(\vec{U}, \vec{u}) \delta_{a \rightarrow i} \delta_{a \rightarrow j} Q_{\hat{\alpha} \rightarrow \hat{a}} Q_{\hat{\beta} \rightarrow \hat{a}} q_{\hat{c} \rightarrow \hat{j}} q_{\hat{d}_5 \rightarrow \hat{j}} q_{\hat{d}_6 \rightarrow \hat{j}} \quad (41)$$

where we used the relation (C4), *i.e.*  $Z_a = Z_{a \rightarrow i} \int Z_i \delta_{a \rightarrow i}$  expressed in terms of field-messages, and we omitted the field dependencies of the  $Q$  and  $q$  second-level messages for brevity. We observe that the field  $u_{a \rightarrow j}$  is integrated over on RHS of (41) and, in addition,  $Z_{a \rightarrow i}$  does not depend on this message (see (C4) for details). We then integrate over this field and make use of the delta function  $\delta_{a \rightarrow j}$ , we get

$$q_{\hat{a} \rightarrow \hat{i}}(\vec{u}) = \frac{\Xi_{\hat{i}}}{\Xi_{\hat{a}}} \int_{\hat{a} \setminus \hat{i}} d\vec{U} d\vec{u} Z_{a \rightarrow i}^y(\vec{U}, \vec{u}) \delta_{a \rightarrow i} \tilde{Q}_{\hat{\alpha} \rightarrow \hat{a}} \tilde{Q}_{\hat{\beta} \rightarrow \hat{a}} \tilde{q}_{\hat{c} \rightarrow \hat{j}} \tilde{q}_{\hat{d}_5 \rightarrow \hat{j}} \tilde{q}_{\hat{d}_6 \rightarrow \hat{j}}. \quad (42)$$

With the tilde notation above we indicated functions  $Q$  and  $q$  with the value of  $u_{a \rightarrow j}$  in their argument replaced by their GBP updates (25) *i.e.* by  $\mathcal{L}_{a \rightarrow j}(\hat{\underline{U}}, \hat{\underline{u}})$ . Equation (42) corresponds to the 1RSB rod-to-vertex Generalized Belief Propagation equation for a 2D lattice model as the EA model and represent the first result of this paper. We recall that the  $Q$  and  $q$  functions above, as their tilde version, are joint probability distributions of several messages therefore equation (42), although analytically consistent, it is very hard, perhaps impossible, to iterate numerically. In Section V we discuss how to take an SP-like ansatz on the  $Q$  and  $q$  functions to simplify the above equation.

### C. Consistency equation for plaquettes and rods

Similarly to the previous section and with an analogous notation, we here want to write more explicitly the consistency equations between plaquettes and rods on the second-level region graph. Equation (37) can be written as

$$\begin{aligned} \frac{1}{\Xi_{\hat{\alpha}}} \int_{\hat{\alpha} \setminus \hat{a}} d\vec{U} d\vec{u} Z_{\alpha}^y(\vec{U}, \vec{u}) \delta_{\hat{p} \rightarrow \hat{r}}^{(4)} \delta_{\hat{p} \rightarrow \hat{r}}^{(v)(8)} \delta_{\hat{r} \rightarrow \hat{v}}^{(8)} \prod^{(4)} Q_{\hat{p} \rightarrow \hat{r}}(\vec{U}, \vec{u}) \prod^{(8)} q_{\hat{r} \rightarrow \hat{v}}(\vec{u}) \\ = \frac{1}{\Xi_{\hat{a}}} Z_a^y(\vec{U}, \vec{u}) \delta_{a \rightarrow i} \delta_{a \rightarrow j} \prod^{(2)} Q_{\hat{p} \rightarrow \hat{r}}(\vec{U}, \vec{u}) \prod^{(6)} q_{\hat{r} \rightarrow \hat{v}}(\vec{u}) \end{aligned} \quad (43)$$

where  $\Xi_{\hat{\alpha}}$  and  $\Xi_{\hat{a}}$  are normalization factors. The second-level messages which appear on the LHS are all those from the parent-to-child region  $\hat{\mu} \rightarrow \nu$  with  $\hat{\mu} \in B(\hat{\alpha})$  and  $\hat{\nu} \in I(\hat{\alpha})$  and the products are over these variables. Pictorially they are represented in Figure 3, right panel, by all the arrows going through the perimeter of the square green region. The  $Q$ 's and  $q$ 's on the RHS are those

going from  $\hat{\mu} \rightarrow \hat{\nu}$  with  $\hat{\mu} \in B(\hat{a})$  and  $\hat{\nu} \in I(\hat{a})$  and are illustrated, in the same aforementioned picture, as the arrows crossing the perimeter of the rod orange region. The product on this side of the equations are on these same variables.

Delta functions in the equation above enforce GBP equations for the field-messages at both the plaquette-to-rod and rod-to-vertex level and, as for (39), are interpreted as hard-core constraints for the second-level region graph model. Their upper script refers to how many delta functions of the type  $\mu \rightarrow \nu$  are included. All those on the LHS of (43) enforce constraints for the field-messages of the region  $I(\alpha)$  whereas those on the RHS enforce constraints for the region  $I(a)$ . Note that since  $I(a) \subset I(\alpha)$  some constraints appear on both sides.

More explicitly, on the LHS four deltas enforce GBP plauquette-to-rod equations (26) for the first level field-messages  $U_{p \rightarrow r}$  as  $\delta_{p \rightarrow r} = \delta[U_{p \rightarrow r} - \mathcal{F}_{p \rightarrow r}(\hat{U}, \hat{u})]$  and eight deltas enforce equations (27) and (28) for the field messages  $u_{p \rightarrow r}^{(v)}$  as  $\delta_{p \rightarrow r}^{(v)} = \delta[u_{p \rightarrow r}^{(v)} - \mathcal{G}_{p \rightarrow r}^{(v)}(\hat{U}, \hat{u})]$ . The remaining eight deltas enforce the GBP link-to-vertex conditions (25) as  $\delta_{r \rightarrow v} = \delta[u_{r \rightarrow v} - \mathcal{L}_{r \rightarrow v}(\hat{U}, \hat{u})]$  and two of them appear also on the RHS. For clarity, referring to Figure 3, left panel, the field messages  $U$  and  $u$  constrained by the delta functions on the LHS of (43) are illustrated as black, red and white arrows in the interior of the yellow square region  $\hat{\alpha}$ . Whereas the field-messages on the RHS of the same equation constrained by the two delta conditions are illustrated as white arrows in the purple rod region  $\hat{a}$  of the same figure.

Note that, as for the rod-to-vertex equation, the introduction of the delta functions would, in principle, involve a determinant term on both sides. The LHS of (43) should be multiplied by  $|\det(\mathbb{1} - \Delta\mathcal{F})|$  where  $\Delta\mathcal{F}$  is the matrix of all the derivatives of the GBP plaquette-to-rod and rod-to-vertex equations respect to all the field-messages in the plaquette. Similarly, the RHS should be multiplied by  $|\det(\mathbb{1} - \Delta\mathcal{L})|$  which is the same determinant, equal to one, encountered in Sec. IV B. It turns out that, after fixing the gauge of the GBP equations as shown by the rest of this section, also  $|\det(\mathbb{1} - \Delta\mathcal{F})|$  is equal to one (see Appendix D) and therefore there exists no correction to the one-step replica symmetric GBP equations because of the determinant, neither for the plaquette-to-rod nor for the rod-to-vertex equations.

Writing equation (43) more explicitly, with labeling referred to Figure 3 we get:

$$\begin{aligned}
& \frac{1}{\Xi_{\hat{\alpha}}} \int_{\hat{\alpha} \setminus \hat{a}} d\vec{U} d\vec{u} Z_{\alpha}^y(\vec{U}, \vec{u}) \underbrace{Q_{\beta \rightarrow a} Q_{\rho \rightarrow b} Q_{\eta \rightarrow d} Q_{\lambda \rightarrow a} q_{d_1 \rightarrow l} q_{d_2 \rightarrow l} q_{d_3 \rightarrow k} q_{d_4 \rightarrow k} q_{d_5 \rightarrow j} q_{d_6 \rightarrow j} q_{d_7 \rightarrow i} q_{d_8 \rightarrow i}}_{\text{LHS}} \\
& \quad \times \delta_{\alpha \rightarrow a}^{(3)} \delta_{\alpha \rightarrow b}^{(3)} \delta_{\alpha \rightarrow c}^{(3)} \delta_{\alpha \rightarrow d}^{(3)} \delta_{a \rightarrow i} \delta_{a \rightarrow j} \delta_{c \rightarrow j} \delta_{c \rightarrow k} \delta_{d \rightarrow k} \delta_{d \rightarrow l} \delta_{b \rightarrow l} \delta_{b \rightarrow i} \\
& = \frac{1}{\Xi_{\hat{a}}} Z_a^y(\vec{U}, \vec{u}) \delta_{a \rightarrow i} \delta_{a \rightarrow j} \underbrace{Q_{\beta \rightarrow a} Q_{\alpha \rightarrow a} q_{c \rightarrow j} q_{d_5 \rightarrow j} q_{d_6 \rightarrow j} q_{d_7 \rightarrow i} q_{d_8 \rightarrow i} q_{b \rightarrow i}}_{\text{RHS}} \quad (44)
\end{aligned}$$

where, on both side of the equation, common probability terms have been underlined and field dependencies of the  $Q$  and  $q$  second-level messages are omitted. To shorten notation, terms like  $\delta_{\alpha \rightarrow a}^{(3)}$  above stands for  $\delta_{\alpha \rightarrow a} \delta_{\alpha \rightarrow a}^{(i)} \delta_{\alpha \rightarrow a}^{(j)}$  and enforce GBP constraints for the fields  $U_{\alpha \rightarrow a}$ ,  $u_{\alpha \rightarrow a}^{(i)}$  and  $u_{\alpha \rightarrow a}^{(j)}$  respectively. Let us note that all the underlined probabilities of field-messages have no dependence on the fields which belong to the region  $\hat{\alpha} \setminus \hat{a}$  but, rather, they only depend on the fields  $\vec{U}$  and  $\vec{u}$  which are in the second-level region  $\hat{a}$ . Using the left panel of Figure 2 as a reference, the underlined second level messages only depend on the field-messages illustrated as black arrows which cross or are internal of the purple rod region  $\hat{a}$ . Therefore since the integration on the LHS of (44) is only over the field-messages in  $\hat{\alpha} \setminus \hat{a}$  (the blue arrows in the aforementioned figure), the underlined  $Q$ 's and  $q$ 's can be simplified on both sides. Using the relation among partition functions of different regions contained in the Appendix, *i.e.* (C5) here written in terms of field messages as  $Z_\alpha = Z_{\alpha \rightarrow a} \int Z_a \delta_{\alpha \rightarrow a} \delta_{\alpha \rightarrow a}^{(i)} \delta_{\alpha \rightarrow a}^{(j)}$ , we can then rewrite (44) as

$$\begin{aligned} \delta_{a \rightarrow i} \delta_{a \rightarrow j} Q_{\alpha \rightarrow a} q_{b \rightarrow i} q_{c \rightarrow j} &= \frac{\Xi_{\hat{a}}}{\Xi_{\hat{\alpha}}} \int_{\hat{\alpha} \setminus \hat{a}} d\vec{U} d\vec{u} Z_{\alpha \rightarrow a}^y(\vec{U}, \vec{u}) Q_{\rho \rightarrow b} Q_{\eta \rightarrow d} Q_{\lambda \rightarrow a} q_{d_1 \rightarrow l} q_{d_2 \rightarrow l} q_{d_3 \rightarrow k} q_{d_4 \rightarrow k} \\ &\times \delta_{\alpha \rightarrow a} \delta_{\alpha \rightarrow a}^{(i)} \delta_{\alpha \rightarrow a}^{(j)} \delta_{\alpha \rightarrow b} \delta_{\alpha \rightarrow b}^{(l)} \delta_{\alpha \rightarrow b}^{(i)} \delta_{\alpha \rightarrow c} \delta_{\alpha \rightarrow c}^{(j)} \delta_{\alpha \rightarrow c}^{(k)} \delta_{\alpha \rightarrow d} \delta_{\alpha \rightarrow d}^{(k)} \delta_{\alpha \rightarrow d}^{(l)} \\ &\times \delta_{a \rightarrow i} \delta_{a \rightarrow j} \delta_{c \rightarrow j} \delta_{c \rightarrow k} \delta_{d \rightarrow k} \delta_{d \rightarrow l} \delta_{b \rightarrow l} \delta_{b \rightarrow i} \end{aligned} \quad (45)$$

Let us note that the two deltas  $\delta_{a \rightarrow i}[u_{a \rightarrow i} - \mathcal{L}_{a \rightarrow i}(\hat{\underline{U}}, \hat{\underline{u}})]$  and  $\delta_{a \rightarrow j}[u_{a \rightarrow j} - \mathcal{L}_{a \rightarrow j}(\hat{\underline{U}}, \hat{\underline{u}})]$  only depend on field-messages in  $\hat{a}$ , *i.e.* on first-level field-messages which are not integrated over in the above equation. Therefore, although  $\hat{a}$  contains fourteen field-messages, equation (45) only holds information for the subset where the two messages  $u_{a \rightarrow i}$  and  $u_{a \rightarrow j}$  in  $I(a)$  are determined by the eight messages in  $B(a)$  through the constraints  $\delta_{a \rightarrow i}$  and  $\delta_{a \rightarrow j}$ . Otherwise the meaning of this aforementioned equation is trivially  $0 = 0$ .

In the following we want to proceed integrating out all the deltas on the RHS of (45) which do not enforce constraints on those fields-messages directed to the region  $I(a)$ . Referring to Figure 2, left panel, we want to integrate out all the fields internal to the yellow square region  $I(\alpha)$  (pictured as blue arrows), except those directed to or internal at the purple regions  $I(a)$  (pictured as black arrows).

First of all we impose gauge fixing conditions for the GBP equations at the plaquette level as discussed in Sec. IIIB and in reference [36]. According to this gauge fixing procedure, at the first-level region graph we set to zero four of the messages  $u_{\alpha \rightarrow r}^{(v)}$  in  $I(r)$ . This means that, at the second-level region graph, four over eight deltas  $\delta_{\alpha \rightarrow r}^{(v)}(u_{\alpha \rightarrow r}^{(v)} - G_{\alpha \rightarrow r}^{(v)}(\vec{U}, \vec{u}))$  are changed to  $\delta(u_{\alpha \rightarrow r}^{(v)})$ . For each rod in the region graph there exist two of these deltas (see for example Fig 2), the choice of which to change accordingly to the fixing gauge conditions discussed in [36] is totally arbitrary,

we then choose to change  $\delta_{\alpha \rightarrow a}^{(i)}, \delta_{\alpha \rightarrow b}^{(l)}, \delta_{\alpha \rightarrow d}^{(k)}, \delta_{\alpha \rightarrow c}^{(j)}$ . Among these deltas, we integrate out those not referred to messages heading to  $a$ , therefore just  $\delta_{\alpha \rightarrow b}^{(l)}, \delta_{\alpha \rightarrow d}^{(k)}, \delta_{\alpha \rightarrow c}^{(j)}$ . The effect of this integration on the rest of the remaining terms is to set to zero the related field-messages wherever else they appear.

We then integrate out three over four of the delta functions which fix the GBP conditions for the fields-messages  $U_{\alpha \rightarrow r}$ , *i.e.* which enforce (26), in particular we integrate out  $\delta_{\alpha \rightarrow b}, \delta_{\alpha \rightarrow c}, \delta_{\alpha \rightarrow d}$ . Let us observe that the function  $\mathcal{F}(\vec{U}, \vec{u})$  which appears in  $\delta_{\alpha \rightarrow r}(U_{\alpha \rightarrow r} - \mathcal{F}_{\alpha \rightarrow r}(\vec{U}, \vec{u}))$  only depends on the external field-messages in  $B(\alpha)$  which, at the GBP level for the plaquette  $\alpha$ , are known values. Therefore the effect of this integration on the remaining terms is to change the value of the  $U_{\alpha \rightarrow r}$  by the values of some external field-messages, wherever they appear.

We continue by integrating out three over four of the remaining  $u_{\alpha \rightarrow r}^{(v)}$  field-messages from plaquette-to-rod of the vertex type and make use of the relative delta functions which enforce the GBP equations (28) for these fields. Explicitly we integrate out  $\delta_{\alpha \rightarrow b}^{(i)}, \delta_{\alpha \rightarrow d}^{(l)}, \delta_{\alpha \rightarrow c}^{(k)}$ . Let us note that the RHS of the equations (28) enforced by these deltas is a function of the external messages, *i.e.* messages in  $B(\alpha)$ , and depend linearly on one message which belong to  $I(\alpha)$ . To write it more explicitly, we could for instance write one of these deltas as  $\delta_{\alpha \rightarrow a}^{(j)}(u_{\alpha \rightarrow a}^{(j)} - (h(\vec{U}_{ex}, \vec{u}_{ex}) - u_{c \rightarrow j}))$  where the label *ex* stands for “externals”, *i.e.* those fields are in  $B(\alpha)$  and  $h(\vec{U}_{ex}, \vec{u}_{ex}) = \mathcal{G}(\vec{U}, \vec{u}) + u_{c \rightarrow j}$  from equation (28). Therefore, the integration over these deltas has the effect of changing messages of the kind  $u_{\alpha \rightarrow a}^{(j)}$  into a function of the external messages and of one internal message of the kind  $u_{c \rightarrow j}$ .

We observe that this effect reflects on the link-to-vertex delta function  $\delta_{b \rightarrow l}, \delta_{d \rightarrow k}, \delta_{c \rightarrow j}$  in a peculiar way. Indeed after these and previous integrations, the GBP functions contained in these deltas become only functions of the external messages, for instance  $\delta_{c \rightarrow j}(u_{c \rightarrow j} - \mathcal{L}_{c \rightarrow j}(\vec{U}_{ex}, \vec{u}_{ex}))$ . This is due to the linear dependence that these latter delta functions have on the internal message of the kind  $u_{c \rightarrow i}$  and it entails simplifications (see equation (B5) for details).

We carry on by integrating out  $\delta_{b \rightarrow l}, \delta_{d \rightarrow k}$  which only bring dependencies on the external messages in the remaining terms. After these latter integration, we are left with the deltas for the fields both heading to and internal at the region  $I(a)$ , *i.e.*  $\delta_{a \rightarrow j}, \delta_{a \rightarrow i}, \delta_{\alpha \rightarrow a}, \delta_{\alpha \rightarrow a}^{(i)}, \delta_{\alpha \rightarrow a}^{(j)}, \delta_{b \rightarrow i}, \delta_{c \rightarrow j}$  and, in addition, with  $\delta_{d \rightarrow l}(u_{d \rightarrow l} - \mathcal{L}_{d \rightarrow l}(\vec{U}_{ex}, \vec{u}_{ex}, u_{c \rightarrow k}))$  and  $\delta_{c \rightarrow k}(u_{c \rightarrow k} - \mathcal{L}_{c \rightarrow k}(\vec{U}_{ex}, \vec{u}_{ex}, u_{a \rightarrow j}))$ . Therefore if we now integrate out  $\delta_{d \rightarrow l}$ , the field  $u_{d \rightarrow l}$  will be replaced by a function of both the external messages and  $u_{c \rightarrow k}$ . We carry on by integrating out also  $\delta_{c \rightarrow k}$ , as a consequence  $u_{c \rightarrow k}$  will be replaced by a function of both the external messages and of  $u_{a \rightarrow j}$ . Equation (45) after all these

integrations reads as:

$$\begin{aligned} \delta_{a \rightarrow i} \delta_{a \rightarrow j} Q_{\alpha \rightarrow a} q_{b \rightarrow i} q_{c \rightarrow j} &= \frac{\Xi_{\hat{a}}}{\Xi_{\hat{a}} \int_{\hat{a} \setminus \hat{a}}} d\vec{U} d\vec{u} Z_{\alpha \rightarrow a}^y(\vec{U}, \vec{u}) \tilde{Q}_{\rho \rightarrow b} \tilde{Q}_{\eta \rightarrow d} \tilde{Q}_{\lambda \rightarrow c} \tilde{q}_{d_1 \rightarrow l} \tilde{q}_{d_2 \rightarrow l} \tilde{q}_{d_3 \rightarrow k} \tilde{q}_{d_4 \rightarrow k} \\ &\times \delta_{\alpha \rightarrow a} \delta_{\alpha \rightarrow a}^{(i)} \delta_{\alpha \rightarrow a}^{(j)} \tilde{\delta}_{b \rightarrow i} \tilde{\delta}_{c \rightarrow j} \delta_{a \rightarrow i} \delta_{a \rightarrow j} \end{aligned} \quad (46)$$

where the tilde notation of the second-level messages and on the deltas means that some of their arguments have been replaced by their GBP updates because of the integration over all the other deltas listed above. Let us observe that both sides of the equation (46) depend on all the field-messages in the region  $\hat{a}$ , *i.e.*  $B(a) \cup I(a)$  which are pictorially represented in Figure 2, left panel, as all the black arrows entering and internal to the purple rod region. Equation (46) corresponds to the 1RSB plaquette-to-rod GBP equation for the 2D EA model. Since the  $Q$ 's and  $q$ 's functions above are joint probability distribution of messages, similar considerations as those in the text below (42) apply here.

## V. A CLASS OF GENERALIZED SURVEY PROPAGATION EQUATIONS

We start by explaining the approach to Survey Propagation (SP) described in [1]. The success of Belief Propagation is to a large extent due to its moving information forward. Indeed, this is why these algorithms are referred to as propagation. SP is a special class of solutions to the general second-level, 1RSB, BP equations which has the same property and for which the fixed points can therefore be found by forward iteration.

In the following we want to verify whether the second-level 1RSB GBP equations presented in the previous section admit a special class of solutions of the SP-type and can therefore be written as a class of Generalized Survey Propagation equations. The special solutions of SP are derived by assuming that a second-level message  $q_{\hat{r} \rightarrow \hat{v}}$  only depends on the one first-level message going the same way, *i.e.* on  $m_{a \rightarrow i}$  [1]. Taking the same assumption, in our case, means that all the  $Q_{\hat{p} \rightarrow \hat{r}}$ 's only depend on the underlining message  $m_{\hat{p} \rightarrow \hat{r}}$  function of two spins and therefore parametrized by three  $u$ -fields, whereas all the  $q_{\hat{r} \rightarrow \hat{v}}$ 's depend on the message  $m_{\hat{r} \rightarrow \hat{v}}$  function of one spin and therefore parametrized by one  $u$ -field. In other words, plaquette-to-rod messages as  $Q_{\hat{\alpha} \rightarrow \hat{a}}$  will be assumed to be function like  $Q_{\hat{\alpha} \rightarrow \hat{a}}(U_{\alpha \rightarrow a}, u_{\alpha \rightarrow a}^{(i)}, u_{\alpha \rightarrow a}^{(j)})$  and rod-to-vertex message to be  $q_{a \rightarrow i}(u_{a \rightarrow i})$ . Following the SP ansatz we assume that this is so for all the second-level messages on RHS of the rod-to-vertex equation (41) and we verify that the assumption is preserved on the LHS. We observe that, the field-message  $u_{a \rightarrow j}$  only appears on RHS as an argument of the delta function  $\delta_{a \rightarrow j}$  which

can be integrated out to give

$$\begin{aligned}
& q_{\hat{a} \rightarrow \hat{i}}(u_{a \rightarrow i}, u_{b \rightarrow i}, u_{d_7 \rightarrow i}, u_{d_8 \rightarrow i}) \\
&= \frac{\Xi_{\hat{i}}}{\Xi_{\hat{a}}} \int_{\hat{a} \setminus \hat{i}} d\vec{U} d\vec{u} Z_{a \rightarrow i}^y(U_{\alpha \rightarrow a}, u_{\alpha \rightarrow a}^{(i)}, u_{\alpha \rightarrow a}^{(j)} U_{\beta \rightarrow a}, u_{\beta \rightarrow a}^{(i)}, u_{\beta \rightarrow a}^{(j)}, u_{c \rightarrow j}, u_{d_5 \rightarrow j}, u_{d_6 \rightarrow j}) \\
&\quad \times \delta_{a \rightarrow i} Q_{\hat{\alpha} \rightarrow \hat{a}} Q_{\hat{\beta} \rightarrow \hat{a}} q_{\hat{c} \rightarrow \hat{j}} q_{\hat{d}_5 \rightarrow \hat{j}} q_{\hat{d}_6 \rightarrow \hat{j}}.
\end{aligned} \tag{47}$$

Above we have written out the arguments of  $q_{\hat{a} \rightarrow \hat{i}}$  (on LHS) and  $Z_{a \rightarrow i}$  (on RHS) explicitly and we recall that the integration turns out to be over all the messages appearing as argument of  $Z_{a \rightarrow i}$ . Compared to (42) we have now gained that the  $Q$  and  $q$  functions appearing on RHS are the original ones which, by assumption, still only depend on one underlying first-level message. Let us now consider on what arguments actually depends  $q_{\hat{a} \rightarrow \hat{i}}$ . Obviously it depends on its first argument since  $u_{a \rightarrow i}$  appears explicitly on RHS in the delta function  $\delta_{a \rightarrow i}(u_{a \rightarrow i} - \mathcal{L}_{a \rightarrow i}(\vec{U}, \vec{u}))$ . However, it does not depend on its other three arguments since neither  $u_{b \rightarrow i}$  nor  $u_{d_7 \rightarrow i}$  nor  $u_{d_8 \rightarrow i}$  appear on RHS. Therefore the assumption that second-level messages only depend on the underlying first-level message going the same way is preserved by the iteration of (47). We observe, in addition, that on a tree topology the above equation reduces to the SP equation for the single instance case [1, 8].

We then consider the plaquette-to-rod equation (46) and take the same kind of SP-ansatz for the second-level messages on its RHS and observe that, as for equation (47), they no longer depend on several external messages brought into their argument by the integration of the internal delta functions performed in the previous section. We therefore remove the tilde notation for them and, after the SP-like ansatz, equation (46) reads:

$$\begin{aligned}
\delta_{a \rightarrow i} \delta_{a \rightarrow j} Q_{\alpha \rightarrow a} q_{b \rightarrow i} q_{c \rightarrow j} &= \frac{\Xi_{\hat{a}}}{\Xi_{\hat{\alpha}}} \int_{\hat{\alpha} \setminus \hat{a}} d\vec{U} d\vec{u} Z_{\alpha \rightarrow a}^y(\vec{U}, \vec{u}) Q_{\rho \rightarrow b} Q_{\eta \rightarrow d} Q_{\lambda \rightarrow a} q_{d_1 \rightarrow l} q_{d_2 \rightarrow l} q_{d_3 \rightarrow k} q_{d_4 \rightarrow k} \\
&\quad \times \delta_{\alpha \rightarrow a} \delta_{\alpha \rightarrow a}^{(i)} \delta_{\alpha \rightarrow a}^{(j)} \tilde{\delta}_{b \rightarrow i} \tilde{\delta}_{c \rightarrow j} \delta_{a \rightarrow i} \delta_{a \rightarrow j}
\end{aligned} \tag{48}$$

To shorten notation we here do not write explicitly all the dependencies of the partition function  $Z_{\alpha \rightarrow a}$  but it turns out that it depends exactly on all the  $U$  and  $u$  messages which appear as argument of the SP-like  $Q$ 's and  $q$ 's on the RHS (see Appendix C) and, as for (47), the integration is precisely on all of them. Using the left panel of Figure 2 as a reference, the integration on the RHS is over all the field-messages represented as blue arrows entering the yellow square region. Let us now look at which are the field dependencies of the RHS of (48). It depends on the fields appearing in the delta functions  $\delta_{\alpha \rightarrow a}, \delta_{\alpha \rightarrow a}^{(i)}, \delta_{\alpha \rightarrow a}^{(j)}, \tilde{\delta}_{b \rightarrow i}$  and  $\tilde{\delta}_{c \rightarrow j}$  which are not integrated over on the RHS. These dependencies are encoded in a SP-like dependence of the  $Q$  and  $q$ 's on the LHS.



In addition, the RHS also depends on all the fields contained in the delta functions  $\delta_{a \rightarrow i}$  and  $\delta_{a \rightarrow j}$  which are also present on the LHS. To give a pictorial representation then, referring to Figure 3, left panel, the delta functions on the RHS depends on the fields represented as red and white arrows. These dependencies are encoded on the LHS by the  $Q$  and the  $q$ 's and the two deltas. The RHS also depends on the fields represented as black arrows entering the purple region in the same figure. These dependencies are encoded on the LHS by the functions  $\mathcal{L}(\vec{U}, \vec{u})$  in the two delta functions  $\delta_{a \rightarrow i}(u_{a \rightarrow i} - \mathcal{L}_{a \rightarrow i}(\vec{U}, \vec{u}))$  and  $\delta_{a \rightarrow j}(u_{a \rightarrow j} - \mathcal{L}_{a \rightarrow j}(\vec{U}, \vec{u}))$ .

Summarizing, equation (47) and (48) remain consistent after the SP-like ansatz is taken and therefore represent a class of generalized GBP equations for the distributions of the field messages, *i.e.*  $Q(U, u, u)$  and  $q(u)$ , similarly as the Survey Propagation [7–10] equations are Belief Propagation equations for the distribution of the field-messages, *i.e.*  $q(u)$ . Although these equations are consistent, they would naturally be cumbersome to use in practice. Indeed the right-hand side is a convolution of many terms and, more importantly, the left-hand side is a product of several distributions and delta functions. Despite the delta functions appearing on the LHS could be integrated on both side of the equations, the product of distributions on the LHS does not allow to use a population dynamic algorithm to iterate the equations. Therefore, the paramagnetic regime, where the equations very much simplify and the product of distribution on the LHS reduces to a single distribution  $Q(u)$ , seems to be the only regime where the equations can be iterated (see [21] for a pioneering contribution about simulations in the paramagnetic phase of the average case).

### A. Comparison with the replica cluster variational method

In an earlier effort Rizzo and co-authors [21] devised a replica cluster variational method (CVM) for studying the EA spin glass model. Here we discuss the relationship between our work and this replica cluster variational approach.

Reference [21] starts applying the CVM formulation to the variational free energy of a  $n$ -times replicated system. Through a standard variational approach replica-GBP equations are derived. The main idea of the authors is then to use a replica symmetric or a more general Parisi's hierarchical ansatz in order to send the number of replicas to zero and so obtain the corresponding RS or (RSB) GBP equations. The core of [21] is dedicated to the analytical and numerical study of the paramagnetic and spin glass replica-symmetric phase of 2D EA model, by considering a plaquette approximation of the averaged free energy. In the Appendices, the authors also show the GBP equations for a generic  $k$ -RSB ansatz, both at the averaged and single instance case.

Our method, on the other hand, starts from the assumption that each GBP fixed point corresponds to a macroscopic state at the 1RSB level. We then build a factor graph model at the 1RSB level in which variable nodes are functions (GBP messages at the RS level) and factor nodes are functional of these GBP messages. The GBP equations on this new factor graph model is our 1RSB GBP scheme. We then focus on the single instance case and therefore do not perform averages over the quenched disorder. Generally speaking, both the replica CVM and our approach can be in principle be used to derive RSB GBP schemes for single instances as well as average cases.

It is interesting to observe that these two approaches provide the same kind of equations for non-redundant region graphs or, even more simply, tree-like topologies. When the graph considered shows redundancy, the sets of equations provided by the two approaches are different and result to be the same only in the paramagnetic phase. We therefore expect, in this regime, a physical behaviour different to the one studied and encountered in [21]. The dissimilarities between the two approaches mainly come from a different derivation. We list them shortly in the following.

In our approach we start assuming that parent-to-child second-level messages depend on all the  $u$ -field contained in the child region. This means that, plaquette-to-rod  $Q$  messages depend on 14 field-messages (see dependencies in eq. (35)) whereas rod-to-vertex  $q$  messages depend on 4 field-messages (see eq. (36)). After the SP ansatz though, our parametrization of the messages turns out to be the same as Rizzo *et al.* (see Sec. V).

Second, in [21] the authors point out that the replica CVM  $Q$ -messages, at the RS level and in the average case, are positive definite only in the paramagnetic phase - where their equations coincide with ours - and therefore can be interpreted as populations only in this regime. As soon as the temperature is lowered below the spin glass transition, they observe that these messages are not necessarily positive definite, therefore they should not be interpreted as probability functionals. In the CVM treatment, then, only the beliefs of second-level regions have a probabilistic meaning. This non-positive definite property indicates that the messages at the RS levels are correlated and redundant. In our formulation, before the SP ansatz is taken as discussed Sec. V, the 1RSB parent-to-child  $Q$ -messages depend on all the  $u$ -fields in the child region, as stated above and in eqs. (35) and (36). This parametrization accounts for correlations among  $u$ -messages and we therefore believe that the  $Q$ 's could be considered as probability functionals. The SP ansatz taken in Section V corresponds to project these functionals on a subspace in order to reduce their dimensionality. We conjecture that the non-positive definite property of the  $Q$ 's would only appear for the average case. In this case all the parent-to-child  $Q$ 's are, indeed, not region dependent and there exist only one type of plaquette-to-rod and rod-to-vertex second-level message. These latter appear in

a convoluted form in the plaquette-to-rod equation and therefore the achievement of a fixed point requires to relax the positive-definite property of the  $Q$ 's. Such extra assumption should not be needed in the single instance case.

On a technical and, more important, physical point of view, we also mention that every  $k$ -RSB formulation of the GBP equations should deal with the gauge fixing of these equations at the first-level. This issue was firstly raised in [36] by the same authors of [21] and following these considerations we addressed this issue in Section III B. The same could be done in the CVM replica scenario but at the time of [21] this was not yet completely understood.

## VI. DISCUSSION

We have considered a one-step replica symmetry breaking description of the Edwards-Anderson model in 2D through a second-level statistical model built on the extremal points of a Kikuchi approximation. These extremal points can be computed as fixed points of Generalized Belief Propagation (GBP) in a parent-to-child message-passing scheme on a region graph with redundancy. We have discussed the fact that these GBP equations exhibit a gauge invariance and how that can be corrected for by setting some parts of the GBP messages to zero. We have then shown that a second-level theory can be constructed where the variables are probability distributions over GBP messages, and the fixed point equations are consistency equations between marginal probabilities of regions in the second-level region graph. We have further shown that this theory has a set of solutions analogous to Survey Propagation (SP) where each second-level message in the second-level model depends only on the corresponding first-level message in the first-level model. For this to be possible it is necessary to again use the gauge invariance and gauge fixing of GBP.

Three comments impose themselves. The first concerns the resulting Generalized Survey Propagation (GSP) equations given in final form as equations (47) and (48). While it would in principle be possible to simulate them numerically, in practice that would be quite a cumbersome task due to their high dimensionality and the product of three messages on the left-hand side of (48). For high enough temperatures, within the paramagnetic phase, the above equations would simplify and numerical treatments become easier although still involving. A numerical investigation of this regime, as reported for instance in [21] for the replica symmetric case, goes beyond the purpose of this paper. Our contribution is therefore mainly of a conceptual nature, pointing to how a generalization of Survey Propagation (SP) to finite-dimensional systems can be carried out, and how complex such an approach appears to have to be.

A second comment concerns the determinant (Jacobian) discussed around equation (33). To our knowledge such determinants have never been discussed before in the literature concerning GSP. It is clear that they must always appear as long as the second-level region graph has loops, and it is also clear that the determinant would be a global characteristic of such a model. On a second-level Bethe lattice, which underlie SP and which has no loops, such determinants is equal to one. We believe that the analysis of such determinants is an interesting though non-trivial task in itself.

Finally, given the prediction that a spin glass is only present at zero temperature in the Edwards-Anderson model in 2D then we should take the zero temperature limit. We note that in the application of 1RSB to random satisfiability problems the temperature is also zero, corresponding to hard constraints, in the first version of SP considered together with the limit of  $y$  equal to zero [8, 9]. All solution clusters or zero-energy states are then weighted equally with no entropic terms depending on the size of each cluster. In a later development it was shown that for the same problems and when  $y$  is larger than zero entropic terms appear that separate clusters of different sizes [1]. It may be that the determinant referred to above complicates the picture in the case at hand, particularly for small  $y$ , and we therefore leave this issue to future work.

**Acknowledgments.** The authors warmly thank Alejandro Lage-Castellanos and Federico Ricci-Tersenghi for valuable discussions which helped us considerably improve upon a previous draft of the paper. Anonymous Referees are also acknowledged for useful comments during the reviewing process of the manuscript. This research is supported by FP7/2007-2013/grant agreement no 290038 (GDF), by the Swedish Science Council through grant 621-2012-2982 (EA), by the Academy of Finland through its Center of Excellence COIN (EA), and by the Natural Science Foundation of China through grant 11225526 (HJZ). GDF and EA thank the hospitality of KITPC and HJZ thanks the hospitality of KTH.

## Appendix A: A derivation of GBP

In this appendix we derive the GBP update equations (16) from a constrained variation of the Kikuchi free energy functional (6). The material is standard and included only for completeness, see [5]. The task is to minimize the Kikuchi free energy functional  $F_K$  in (6) *i.e.*

$$F_K = \sum_{\alpha} c_{\alpha} \sum_{\underline{\sigma}_{\alpha}} P_{\alpha}(\underline{\sigma}_{\alpha}) \left[ E_{\alpha}(\underline{\sigma}_{\alpha}) + \beta^{-1} \log P_{\alpha}(\underline{\sigma}_{\alpha}) \right]$$

where we have introduced the energy of a region as

$$E_\alpha(\underline{\sigma}_\alpha) = \sum_{i \in \alpha} h_i \sigma_i + \sum_{a \in \alpha} J_{ij}^{(a)} \sigma_i \sigma_j.$$

The variation is of the marginal probabilities  $P_\alpha(\underline{\sigma}_\alpha)$  under the constraints that they have to be consistent with each other. This means, in the situation under consideration, that if a rod region  $a$  is a child of plaquette region  $\alpha$  then  $\sum_{\underline{\sigma}_\alpha \setminus \underline{\sigma}_a} P_\alpha(\underline{\sigma}_\alpha) = P_a(\underline{\sigma}_a)$  where  $\underline{\sigma}_\alpha \setminus \underline{\sigma}_a$  is the exclusion set consisting of all spins in  $\alpha$  that are not in  $a$ . Similarly, if a vertex region  $i$  is a child of rod region  $a$  then  $\sum_{\underline{\sigma}_a \setminus \sigma_i} P_a(\underline{\sigma}_a) = P_i(\sigma_i)$ . Introducing Lagrange multipliers  $\lambda_{\alpha|a}(\sigma_l, \sigma_k)$  and  $\lambda_{a|i}(\sigma_i)$  to represent all these constraints, where the spin indexing follows Fig. 1, as well Lagrange multipliers  $q_\alpha$ ,  $q_a$  and  $q_i$  for the normalizations of  $P_\alpha$ ,  $P_a$  and  $P_i$ , we arrive at

$$\begin{aligned} P_i &\propto e^{-\beta E_i} e^{\frac{1}{c_i} \beta \sum_{a:a \rightarrow i} \lambda_{a|i}} \\ P_a &\propto e^{-\beta E_a} e^{-\frac{1}{c_a} \beta \sum_{i:a \rightarrow i} \lambda_{a|i}} e^{\frac{1}{c_a} \beta \sum_{\alpha:\alpha \rightarrow a} \lambda_{\alpha|a}} \\ P_\alpha &\propto e^{-\beta E_\alpha} e^{-\frac{1}{c_\alpha} \beta \sum_{a:\alpha \rightarrow a} \lambda_{\alpha|a}} \end{aligned} \quad (\text{A1})$$

Referring to Fig. 1 for the labeling of the regions, where *e.g.* vertex  $i$  has four rod parents, of which one is  $a$  and the one opposite is called  $d_8$ , and a rod has two plaquette parents, of which one is  $\alpha$  and the other one is  $\beta$ , we can introduce auxiliary normalized quantities

$$m_{a \rightarrow i} \propto e^{\beta \lambda_{d_8|i}} \quad (\text{A2})$$

$$m_{\beta \rightarrow a} \propto e^{-\beta \lambda_{\alpha|a} - \lambda_{d_6|j} - \lambda_{d_7|i}} \quad (\text{A3})$$

In terms of the  $m_{a \rightarrow i}$  and the  $m_{\alpha \rightarrow a}$  the marginal probabilities of the regions then take the GBP output equation form *i.e.*

$$\begin{aligned} P_i &\propto e^{-\beta E_i} \prod_{a:a \rightarrow i} m_{a \rightarrow i} \\ P_a &\propto e^{-\beta E_a} m_{\alpha \rightarrow a} m_{\beta \rightarrow a} \prod_{b:b \rightarrow i \setminus a} m_{b \rightarrow i} \prod_{c:c \rightarrow j \setminus a} m_{c \rightarrow j} \\ P_\alpha &\propto e^{-\beta E_\alpha} m_{\beta \rightarrow a} m_{\rho \rightarrow b} m_{\eta \rightarrow d} m_{\lambda \rightarrow c} \prod_{k=1}^8 m_{d_k \rightarrow}. \end{aligned} \quad (\text{A4})$$

where the dummy argument in  $m_{d_k \rightarrow}$  indicates the vertex in the intersection of  $d_k$  and  $\alpha$ . Varying the messages in (A4) gives the GBP update equations as described in the main text. Alternatively one can check that the marginalization constraints expressed in the Lagrange multipliers themselves, which are

$$\begin{aligned} e^{-\beta \sum_{b:b \rightarrow i \setminus a} \lambda_{b|i}} &\propto \sum_{\sigma_j} e^{-\beta(E_a - E_i) + \beta \lambda_{a|j} - \beta(\lambda_{\alpha|a} + \lambda_{\beta|a})} \\ e^{\beta(\lambda_{a|i} + \lambda_{a|j}) - \beta \lambda_{\beta|a}} &\propto \sum_{\sigma_l, \sigma_k} e^{-\beta(E_\alpha - E_a) - \beta \sum_{c:\alpha \rightarrow c \setminus a} \lambda_{\alpha|c}} \end{aligned} \quad (\text{A5})$$

are equivalent to the GBP update equations for the messages. We observe that nothing in the above derivation really depends on how many variables (spins) there are in the various regions and what their interactions are; the derivation therefore also holds for the second-level GBP.

## Appendix B: GBP equations for the 2D Edwards-Anderson model: from messages to fields.

In this appendix, by parametrizing the plaquette-to-rod and rod-to-vertex messages in terms of cavity fields, we want to rewrite the GBP equations (21) and (23) in terms of fields. As already shown in the main text, a message  $m_{a \rightarrow i}(\sigma_i)$ , which is a probability distribution on the spin  $\sigma_i$  up to a normalization, can be written as  $e^{\beta u_{a \rightarrow i} \sigma_i}$  where  $u_{a \rightarrow i}$  is the cavity field on spin  $i$  from interaction  $a$ . A plaquette-to-rod message  $m_{\alpha \rightarrow a}(\sigma_i, \sigma_j)$  is a probability distribution on two spins and can be written, up to a normalization, as  $e^{\beta(u_{\alpha \rightarrow a}^{(i)} \sigma_i + u_{\alpha \rightarrow a}^{(j)} \sigma_j + U_{\alpha \rightarrow a} \sigma_i \sigma_j)}$  where the new quantity  $U_{\alpha \rightarrow a}$  parametrizes two body interactions. The GBP equations (21) and (23) are then relations between parameters of the types  $u_{a \rightarrow i}$ ,  $u_{\alpha \rightarrow a}^{(i)}$  and  $U_{\alpha \rightarrow a}$ . In addition we remind that the potential functions  $\psi$  present in the GBP equations are given accordingly to the definition in the EA model as in (4), namely  $\psi_i(\sigma_i) = e^{\beta h_i \sigma_i}$  and  $\psi_a(\sigma_{\partial a}) = e^{\beta J_{ij}^{(a)} \sigma_i \sigma_j}$ .

**Rod-to-vertex equation:** To write this equation explicitly we note that the RHS of (23) can be parametrized as a function of the spin  $\sigma_i$  and  $\sigma_j$ :

$$m_{a \rightarrow i}(\sigma_i) = \frac{1}{Z_{a \rightarrow i}} \exp[\beta u_{a \rightarrow i} \sigma_i] = \frac{1}{Z_{a \rightarrow i}} \sum_{\sigma_j} e^{\beta(x_i \sigma_i + x_j \sigma_j + X_{ij} \sigma_i \sigma_j)}, \quad (\text{B1})$$

where

$$x_i = u_{\alpha \rightarrow a}^{(i)} + u_{\beta \rightarrow a}^{(i)}, \quad (\text{B2})$$

$$x_j = h_j + u_{\alpha \rightarrow a}^{(j)} + u_{\beta \rightarrow a}^{(j)} + u_{c \rightarrow j} + u_{d_5 \rightarrow j} + u_{d_6 \rightarrow j}, \quad (\text{B3})$$

$$X_{ij} = J_{ij}^{(a)} + U_{\alpha \rightarrow a} + U_{\beta \rightarrow a}. \quad (\text{B4})$$

Computing the rate  $m_{a \rightarrow i}(1)/m_{a \rightarrow i}(-1)$  with the parametrization given above and taking the logarithm on both sides of the resulting equation, we obtain the expression for the rod-to-vertex field [36]:

$$u_{a \rightarrow i} = x_i + \frac{1}{2\beta} \ln \left[ \frac{\cosh[\beta(x_j + X_{ij})]}{\cosh[\beta(x_j - X_{ij})]} \right] \quad (\text{B5})$$

**Plaquette-to-rod equation:** In the following we will not keep track of the normalization constants of the various messages. With the parametrization of messages in terms of fields given above, equation (21) can be written as:

$$\exp \left[ \beta \left( (u_{\alpha \rightarrow a}^{(i)} + u_{b \rightarrow i}) \sigma_i + (u_{\alpha \rightarrow a}^{(j)} + u_{c \rightarrow j}) \sigma_j + U_{\alpha \rightarrow a} \sigma_i \sigma_j \right) \right] = e^{u_{\rho \rightarrow b}^{(i)} \sigma_i + u_{\lambda \rightarrow c}^{(j)} \sigma_j} K(\sigma_i, \sigma_j) \quad (\text{B6})$$

where the function  $K(\sigma_i, \sigma_j)$ , according to the RHS of (21), reads as:

$$\begin{aligned}
K(\sigma_i, \sigma_j) = \sum_{\sigma_k, \sigma_l} \exp \Big[ & \left( (h_l + u_{\rho \rightarrow b}^{(l)} + u_{d_1 \rightarrow l} + u_{d_2 \rightarrow l} + u_{\eta \rightarrow d}^{(l)}) \sigma_l \right. \\
& (h_k + u_{\eta \rightarrow d}^{(k)} + u_{d_3 \rightarrow k} + u_{d_4 \rightarrow k} + u_{\lambda \rightarrow c}^{(k)}) \sigma_k + \\
& \left. (J_{li}^{(b)} + U_{\rho \rightarrow b}) \sigma_l \sigma_i + (J_{lk}^{(d)} + U_{\eta \rightarrow d}) \sigma_l \sigma_k + (J_{kj}^{(c)} + U_{\lambda \rightarrow c}) \sigma_k \sigma_j \right) \Big].
\end{aligned} \tag{B7}$$

After some simple algebra the plaquette-to-rod fields on the LHS of (B6) are then given by

$$u_{\alpha \rightarrow a}^{(i)} = u_{\rho \rightarrow b}^{(i)} - u_{b \rightarrow i} + \frac{1}{4\beta} \log \left[ \frac{K(1, 1)K(1, -1)}{K(-1, 1)K(-1, -1)} \right], \tag{B8}$$

$$u_{\alpha \rightarrow a}^{(j)} = u_{\lambda \rightarrow c}^{(j)} - u_{c \rightarrow j} + \frac{1}{4\beta} \log \left[ \frac{K(1, 1)K(-1, 1)}{K(1, -1)K(-1, -1)} \right], \tag{B9}$$

$$U_{\alpha \rightarrow a} = \frac{1}{4\beta} \log \left[ \frac{K(1, 1)K(-1, -1)}{K(1, -1)K(-1, 1)} \right]. \tag{B10}$$

where we used a similar notation as in [36] to facilitate understanding.

### Appendix C: Useful relations relative to the the partition functions

In this appendix we want to show some relations among the different partition functions encountered in the main text which are useful for the presentation of the theory in Sec IV. We wish to show that the partition functions of the plaquette, rod and vertex regions are related among each other through the partition functions which appear as normalizing constant in the GBP equations (21) and (23). Let us write explicitly the partition functions for the region considered through the paper:

$$\begin{aligned}
Z_\alpha = & \sum_{\sigma_i, \sigma_j, \sigma_k, \sigma_l} \psi_i(\sigma_i) \psi_j(\sigma_j) \psi_k(\sigma_k) \psi_l(\sigma_l) \psi_a(\sigma_i, \sigma_j) \psi_b(\sigma_l, \sigma_i) \psi_c(\sigma_j, \sigma_k) \psi_d(\sigma_k, \sigma_l) \\
& \times m_{\beta \rightarrow a}(\sigma_i, \sigma_j) m_{\rho \rightarrow b}(\sigma_l, \sigma_i) m_{\eta \rightarrow d}(\sigma_k, \sigma_l) m_{\lambda \rightarrow c}(\sigma_j, \sigma_k) \\
& \times m_{d_1 \rightarrow l}(\sigma_l) m_{d_2 \rightarrow l}(\sigma_l) m_{d_3 \rightarrow k}(\sigma_k) m_{d_4 \rightarrow k}(\sigma_k) m_{d_5 \rightarrow j}(\sigma_j) m_{d_6 \rightarrow j}(\sigma_j) m_{d_7 \rightarrow i}(\sigma_i) m_{d_8 \rightarrow i}(\sigma_i)
\end{aligned} \tag{C1}$$

$$\begin{aligned}
Z_a = & \sum_{\sigma_i, \sigma_j} \psi_i(\sigma_i) \psi_j(\sigma_j) \psi_a(\sigma_i, \sigma_j) m_{\alpha \rightarrow a}(\sigma_i, \sigma_j) m_{\beta \rightarrow a}(\sigma_i, \sigma_j) \\
& \times m_{b \rightarrow i}(\sigma_i) m_{d_8 \rightarrow i}(\sigma_i) m_{d_7 \rightarrow i}(\sigma_i) m_{c \rightarrow j}(\sigma_j) m_{d_5 \rightarrow j}(\sigma_j) m_{d_6 \rightarrow j}(\sigma_j)
\end{aligned} \tag{C2}$$

$$Z_i = \sum_{\sigma_i} \psi_i(\sigma_i) m_{a \rightarrow i}(\sigma_i) m_{b \rightarrow i}(\sigma_i) m_{d_8 \rightarrow i}(\sigma_i) m_{d_7 \rightarrow i}(\sigma_i) \tag{C3}$$

Let us first consider the relation between the rod and vertex partition function. The rod partition function (C2) can be rewritten as

$$\begin{aligned}
Z_a &= Z_{a \rightarrow i} \int \sum_{\sigma_i} \psi_i m_{a \rightarrow i} m_{b \rightarrow i} m_{d_8 \rightarrow i} m_{d_7 \rightarrow i} \delta(m_{a \rightarrow i} - \mathcal{L}_{a \rightarrow i}) dm_{a \rightarrow i} \\
&= Z_{a \rightarrow i} [m_{\alpha \rightarrow a}, m_{\beta \rightarrow a} m_{c \rightarrow j}, m_{d_5 \rightarrow j}, m_{d_6 \rightarrow j}] \\
&\times \int Z_i(m_{a \rightarrow i}, m_{b \rightarrow i} m_{d_8 \rightarrow i} m_{d_7 \rightarrow i}) \delta(m_{a \rightarrow i} - \mathcal{L}_{a \rightarrow i}) dm_{a \rightarrow i}
\end{aligned} \tag{C4}$$

where in the first equality we used the RHS of the GBP equation (23) to recognize  $m_{a \rightarrow i}$  in (C2) and we introduced the delta function to enforce such a GBP equation. The function  $Z_{a \rightarrow i}$  is the constant appearing in equation (23), its and the  $Z_i$ 's dependencies on the messages are written explicitly on the RHS for clarity. Using the expression (C1) and (C2), a similar calculation shows that the plaquette partition function  $Z_\alpha$  can be written as

$$Z_\alpha = Z_{\alpha \rightarrow a} [m_{\rho \rightarrow b}, m_{\eta \rightarrow d}, m_{\lambda \rightarrow c}, m_{d_1 \rightarrow j}, m_{d_2 \rightarrow j}, m_{d_3 \rightarrow k}, m_{d_4 \rightarrow k}] \int Z_a \delta(m_{\alpha \rightarrow a} - \mathcal{F}_{\alpha \rightarrow a}) dm_{\alpha \rightarrow a} \tag{C5}$$

where we used the the RHS of the GBP equation (21) to recognize and substitute terms equivalent to  $m_{\alpha \rightarrow a}$  in (C1) and the delta function to enforce such a GBP equation. The function  $Z_{\alpha \rightarrow a}$  above corresponds to the constant factor of equation (21) with its own dependencies written explicitly for clarity.

#### Appendix D: Computation of determinants

As already pointed out in the main text, a proper formulation of the second-level GBP equations should take into account determinant factors which act as entropic corrections in the re-weighting terms included in the partition functions at the second-level.

We encountered two classes of these corrections, the first one is the determinant appearing in equation (33). This term corresponds to a large  $L \times L$  matrix where  $L$  is the number of GBP messages (plaquette-to-link and link-to-vertex) appearing in the whole square lattice considered. Its contribution can be written as  $|\det(\mathbb{1} - \mathbb{J})|$  where  $\mathbb{J}$  is the  $L \times L$  matrix of all the derivatives of the GBP equations respect to all the plaquette-to-link and link-to-vertex field messages computed at the fixed point. Its explicit computation is a cumbersome task therefore, at a first level of modelling, we neglected this contribution.

We can instead compute explicitly the second class of determinant corrections arising in the main text, which are those appearing in the consistency equations (39) and (43), used to derive



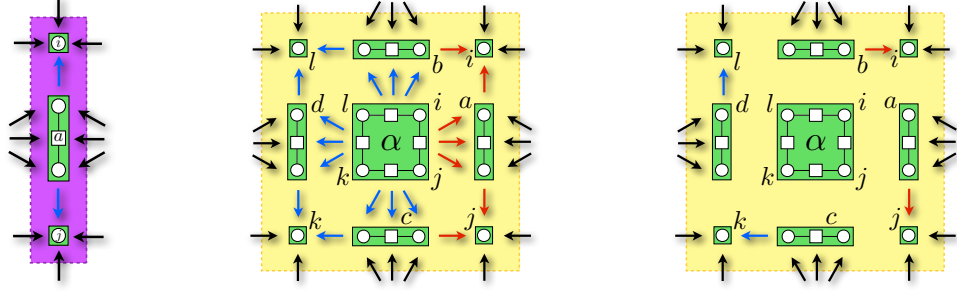


FIG. 4: Left and central panel: second level rod and plaquette region respectively with external field-messages depicted as black arrows and internal depicted either as blue or red ones. Right panel: plaquette region after some of the internal field-messages have been integrated out. All the red arrows in the central and right panel represent fields with gauge fixed (see Section III B and IV C).

the GBP equations at the second-level (see lines below these equations). In the main text we claim that these two determinants are equal to one. Hereby, rather than computing explicitly the matrix of all the derivatives in the two cases and calculate it at the fixed point, we give a simple argument which proves our claim.

Let us recall that determinants appear in the equations because of the integration of the delta functions, we here show that, these integrations can be carried out completely without entering in loops and, therefore, the determinant contribution is equal to one. Hereafter we refer to figure 4 for a pictorial representation. For the case of the equation (39), the term  $|\det(1 - \Delta\mathcal{L})|$  rises because of the presence of the two delta functions  $\delta_{a \rightarrow i}$  and  $\delta_{a \rightarrow j}$ . The fields enforced by these deltas are represented in figure 4, left panel, by blue arrows which (according to their GBP equations) only depends on the external fields of the link region (depicted as black arrows). Therefore the integration over  $u_{a \rightarrow i}$  and  $u_{a \rightarrow j}$  can be carried out and the result will be just the change of the value of these two fields with external ones and hence the determinant contribution turns out to be one.

More involved is the same proof for the term  $|\det(1 - \Delta\mathcal{F})|$  appearing in equation (43). This determinant emerges because of the integration over all the delta functions enforcing GBP conditions for the fields in  $I(\alpha)$ , depicted with (non-black) arrows in figure 4, central panel.

We here want to stress that, to finally prove that also this term is equal to one, it is crucial to fix the gauge (see Section III B and IV C), *i.e.* to change some deltas from  $\delta(u_{x \rightarrow y} - F_{x \rightarrow y}(\dots))$  to  $\delta(u_{x \rightarrow y})$ . This indeed breaks the loop between the GBP equations inside the plaquette region, as it will be clear in the following. Fields which have been changed accordingly to the gauge fixing

are depicted as red arrows in the figure above for clearness.

The evidence is based on a similar procedure encountered in Section IV C where we integrate some delta functions to simplify the equation therein. We start integrating out all the fields  $U_{\alpha \rightarrow a}, U_{\alpha \rightarrow b}, U_{\alpha \rightarrow c}, U_{\alpha \rightarrow d}$  which are the central arrows from the plaquette-to-link in Fig. 4, central panel, noting that, according to their GBP equations, these fields depend only on other external fields (black arrows). We then integrate out all the (right oriented)-fields of the type  $u_{\alpha \rightarrow a}^{(j)}$  (namely  $u_{\alpha \rightarrow a}^{(j)}, u_{\alpha \rightarrow b}^{(i)}, u_{\alpha \rightarrow c}^{(k)}, u_{\alpha \rightarrow d}^{(l)}$ ) and observe that, as mentioned in the main text, these integrations (see equation (B9)) have the final effect of changing the dependence of the (counter-clockwise oriented) fields like  $u_{a \rightarrow i}$  on only external fields (see equation (B5)). We then can carry out the integration over these latter fields, namely  $u_{a \rightarrow i}, u_{b \rightarrow l}, u_{d \rightarrow k}, u_{c \rightarrow j}$ . Once this is done, the remaining fields  $u_{\alpha \rightarrow a}^{(i)}, u_{\alpha \rightarrow b}^{(l)}, u_{\alpha \rightarrow c}^{(j)}, u_{\alpha \rightarrow d}^{(k)}$  also only depends on the external other fields (see equation (B8) considering that messages of the type  $u_{b \rightarrow i}$  have been integrated out and carry only a dependence on the external fields).

We are then left with a picture illustrated in Figure 4, right panel, where there remain only four internal fields to be integrated out. As we can see, in principle, these fields are in loop chain dependence according to their GBP equations. Indeed, using  $\rightarrow$  as a symbol of dependence, we have that  $u_{b \rightarrow i} \rightarrow u_{d \rightarrow l} \rightarrow u_{c \rightarrow k} \rightarrow u_{a \rightarrow j} \rightarrow u_{b \rightarrow i}$  and so on. It is therefore here that, the gauge fixing condition becomes vital to achieve the result. Indeed, after the gauge is fixed, the fields depicted as red arrows in figure 4, right panel, do not share any dependence on other fields and therefore can be integrated out. The field  $u_{c \rightarrow k}$  hence only depends on external fields and can also be integrated out, leaving  $u_{d \rightarrow l}$  depending only on external messages and therefore its integration can also be carried out with no consequences.

This proves that the term  $|\det(\mathbb{1} - \Delta\mathcal{F})|$  is also equal to one as  $|\det(\mathbb{1} - \Delta\mathcal{L})|$  and therefore there are no determinant corrections to the GBP equations on a second-level. A similar argument to the one used here can be applied to show that also the determinant which appear in the SP equation in [1, 2] is actually equal to one. Indeed on tree topologies loops are absent and the integration over the delta functions can be carried out with no side-effects.

- 
- [1] Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.
  - [2] Francesco Zamponi. Mean field theory of spin glasses. *arXiv preprint arXiv:1008.4844*, 2010.
  - [3] Rémi Monasson. Structural glass transition and the entropy of the metastable states. *Physical review*

- letters*, 75(15):2847, 1995.
- [4] M. Mezard, G. Parisi, and M.A. Virasoro. *Spin Glass Theory and Beyond*. World Scientific, Singapore, 1987.
  - [5] Jonathan S Yedidia, William T Freeman, and Yair Weiss. Understanding belief propagation and its generalizations. In *Exploring artificial intelligence in the new millennium*, pages 239–269. Morgan Kaufmann Publishers Inc., 2003.
  - [6] Martin J. Wainwright and Michael I. Jordan. Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*, 1:1–305, 2008.
  - [7] Marc Mézard and Giorgio Parisi. The Bethe lattice spin glass revisited. *Eur. Phys. J. B*, 20(2):217–233, 2001.
  - [8] Marc Mézard, Giorgio Parisi, and Riccardo Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812–815, 2002.
  - [9] Marc Mézard and Riccardo Zecchina. Random k-satisfiability problem: From an analytic solution to an efficient algorithm. *Physical Review E*, 66(5):056126, 2002.
  - [10] Alfredo Braunstein, Marc Mézard, and Riccardo Zecchina. Survey propagation: An algorithm for satisfiability. *Random Structures & Algorithms*, 27(2):201–226, 2005.
  - [11] Jonathan S Yedidia, William T Freeman, Yair Weiss, et al. Generalized belief propagation. In *NIPS*, volume 13, pages 689–695, 2000.
  - [12] Daniel S. Fisher and David A. Huse. Equilibrium behavior of the spin-glass ordered phase. *Phys. Rev. B*, 38:386–411, Jul 1988.
  - [13] Alexander K. Hartmann. Evidence for nontrivial ground-state structure of  $3d \pm j$  spin glasses. *Europhys. Lett.*, 40, 1997.
  - [14] A. Alan Middleton. Numerical investigation of the thermodynamic limit for ground states in models with quenched disorder. *Phys. Rev. Lett.*, 83:1672–1675, Aug 1999.
  - [15] Giorgio Parisi. Some considerations of finite-dimensional spin glasses. *J. Phys. A: Math. Theor.*, 41, 2008.
  - [16] Ryoichi Kikuchi. A theory of cooperative phenomena. *Physical review*, 81(6):988, 1951.
  - [17] Guozhong An. A note on the cluster variation method. *Journal of Statistical Physics*, 52(3-4):727–734, 1988.
  - [18] Alessandro Pelizzola. Cluster variation method in statistical physics and probabilistic graphical models. *Journal of Physics A: Mathematical and General*, 38(33):R309, 2005.
  - [19] Jonathan S Yedidia, William T Freeman, and Yair Weiss. Bethe free energy, kikuchi approximations, and belief propagation algorithms. *Advances in neural information processing systems*, 13, 2001.
  - [20] Jonathan S Yedidia, William T Freeman, and Yair Weiss. Constructing free-energy approximations and generalized belief propagation algorithms. *Information Theory, IEEE Transactions on*, 51(7):2282–2312, 2005.
  - [21] Tommaso Rizzo, Alejandro Lage-Castellanos, Roberto Mulet, and Federico Ricci-Tersenghi. Replica

- cluster variational method. *Journal of Statistical Physics*, 139(3):375–416, 2010.
- [22] A. Lage-Castellanos, R. Mulet, F. Ricci-Tersenghi, and T. Rizzo. Replica cluster variational method: the replica symmetric solution for the 2d random bond ising model. *J. Phys. A: Math. Theor.*, 46, 2013.
  - [23] Hai-Jun Zhou, Chuang Wang, Jing-Qing Xiao, and Zedong Bi. Partition function expansion on region graphs and message-passing equations. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(12):L12001, 2011.
  - [24] Hai-Jun Zhou and Chuang Wang. Region graph partition function expansion and approximate free energy landscapes: Theory and some numerical results. *Journal of Statistical Physics*, 148(3):513–547, 2012.
  - [25] Chuang Wang and Hai-Jun Zhou. Simplifying generalized belief propagation on redundant region graphs. *Journal of Physics: Conference Series*, vol. 473, no. 1, p. 012004. IOP Publishing, 2013.
  - [26] Samuel Frederick Edwards and Phil W Anderson. Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5(5):965, 1975.
  - [27] Ingo Morgenstern and K Binder. Magnetic correlations in two-dimensional spin-glasses. *Physical Review B*, 22(1):288, 1980.
  - [28] Lawrence Saul and Mehran Kardar. Exact integer algorithm for the two-dimensional  $\pm j$  ising spin glass. *Physical Review E*, 48(5):R3221, 1993.
  - [29] T Jörg, J Lukic, E Marinari, and OC Martin. Strong universality and algebraic scaling in two-dimensional ising spin glasses. *Physical review letters*, 96(23):237205, 2006.
  - [30] Creighton K Thomas, David A Huse, and A Alan Middleton. Zero-and low-temperature behavior of the two-dimensional  $\pm j$  ising spin glass. *Physical review letters*, 107(4):047203, 2011.
  - [31] Creighton K Thomas and A Alan Middleton. Exact algorithm for sampling the two-dimensional ising spin glass. *Physical Review E*, 80(4):046708, 2009.
  - [32] Michael Chertkov and Vladimir Y Chernyak. Loop series for discrete statistical models on graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(06):P06009, 2006.
  - [33] Vladimir Y Chernyak and Michael Chertkov. Loop calculus and belief propagation for q-ary alphabet: Loop tower. In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pages 316–320. IEEE, 2007.
  - [34] Andrea Montanari and Tommaso Rizzo. How to compute loop corrections to the bethe approximation. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(10):P10011, 2005.
  - [35] Ryuhei Mori. Loop calculus for nonbinary alphabets using concepts from information geometry. *Information Theory, IEEE Transactions on*, 61(4):1887–1904, 2015.
  - [36] Eduardo Domínguez, Alejandro Lage-Castellanos, Roberto Mulet, Federico Ricci-Tersenghi, and Tommaso Rizzo. Characterizing and improving generalized belief propagation algorithms on the 2d edwards–anderson model. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(12):P12007, 2011.
  - [37] A Lage-Castellanos, R Mulet, and F Ricci-Tersenghi. Message passing and monte carlo algorithms:

Connecting fixed points with metastable states. *EPL (Europhysics Letters)*, 107(5):57011, 2014.